Three Essays in Applied Economics

Eugenia Suarez

A thesis submitted for the degree of Doctor of Philosophy

Department of Economics

University of Essex

January 2016

To Ricardo

Acknowledgements

Finishing the PhD is the conclusion to a long journey that has been everything but uneventful where many people have been involved.

Special mention goes to my academic supervisors, Giovanni Mastrobuoni and George Symeonides, for their guidance, and support throughout this process. I would like to thank my sponsor CONACYT whose support through their scholarship program allowed me to complete my studies.

The PhD is a very difficult experience, but my friends and colleagues, Debora Luna, Tanya Zenina, Miguel Sarzosa and Steven Pennings helped throughout this journey to make it more achievable. Personally and professionally, I owe a big thank you to Jenny Guardado. She has actively supported me to pursue the PhD.

I am also grateful to my parents Rosa Maria and Joaquin, my sister Ana, my brothers Federico and Santiago, along with my grandparents that were always there to support me. They were always kind with their words and motivation.

Finally, I could have never reached the goal of becoming a PhD if it was not for my husband, Ricardo. I find it difficult to express my appreciation because it is simply immeasurable. He is my rock. I am grateful to my husband not just because he has given up so much and endured the difficult times the past years brought. He has shared this entire journey with me, always offering unbounded love, support, and encouragement. I dedicate this thesis to him.

Contents

Summary

In this thesis I present three essays that explore various economic situations on strategic choices from different perspectives: the individuals' strategic decision to work on the informal/formal sector, the US strategic decision on the provision of foreign aid, and the firm's strategic decision to engage in Corporate Social Responsibility (CSR).

The first essay presents an analysis on the effect of income taxes and its effect on worker's transitions towards informality. We find that an increase in average tax rate leads to a statistically significant increase in transitions towards informality for women and those with low incomes.

The second essay offers evidence of how patterns of US foreign aid to Latin America differ from aid allocation observed elsewhere. We find that while political institutions and events in recipient countries greatly influence US aid allocations, the *ideological orientation* of US administrations can explain part of the divergent patters of aid towards Latin America.

Finally, the third essay studies two possible mechanisms that affect the decision of a firm to engage in CSR: the role of growth in value added and workers' preferences. The results suggest that firms engage in CSR in times of economic prosperity; peer effects are increasingly important in a firm's decision to engage in CSR when the proportion of firms within an industry increases. And finally, I find a weak link between workers' preferences and a firm's decision to engage in CSR activities related with diversity.

Chapter 1

Informality and Taxes in Mexico

Abstract

A large literature has argued that taxes on formal sector labor encourage high rates of informal employment in developing and emerging countries. Despite this, there are few estimates of the strength of this channel, and how it varies across segments of the population. This paper seeks to fill that gap by exploiting plausibly exogenous changes on the tax schedule in Mexico over 2005-14 which changed the average tax rate faced by some workers more than others. We use large rotating panel survey, which allows us to control for individual effects, time effects, seasonality and a range of demographic, regional and industry-level factors. As tax rates depend on income, which is potentially endogenous to informal status, we follow Auten and Carroll (1999) and instrument actual average tax rates with the change in tax rates based on lagged income. We find that an increase in average tax rate leads to a statistically significant increase in transitions towards informality for women and those with low incomes. The results are driven by low-income women, for whom a 1 percentage point in average labor tax rates increases the probability of being employed in the informal sector work by 1.5 percentage points. The result is even stronger for women in rural areas. In contrast, there is little evidence that higher tax rates increase the probability of transition to informal sector work for men or those on higher income.

[0]Co-author: Steven Pennings (email: spennings@worldbank.org)

1.1 Introduction

Informal employment is a characteristic of the labor market in most developing economies. In Latin America[1] it accounts for around 50 percent of total employment (Gasparini and Tornarolli, 2007), with a similar percentage for Mexico (Maloney, 2004; INEGI, 2014) (see Figure 1.1).

Reducing informality is a pressing issue for policymakers all over the world and Mexico is not an exception. Many scholars argue that the persistence of informality in recent years in Mexico has been a major reason for its low productivity growth (Matias Busso, 2012; Jorge Alonso Ortiz, 2011; Ordoñez, 2014; Levy, 2008). To reduce informality, we need to firstly understand what are its determinants. The literature suggests that income and social security taxes, government effectiveness, the stringency of regulations and the economic environment are possible candidates (Schneider, Buehn, and Montenegro, 2010; Dougherty and Escobar, 2013).

The focus of this paper is on the effect of taxes and social security contributions on transitions to and away from the informal labor sector, which are thought by many to be among the main causes of informality (Levy, 2008; Pedro S. Amaral, 2006; Mariano Bosch and Campos-Vazquez, 2014). We focus on a labor market definition of informal employment based on a lack of registration at the Mexican Social Security Institute (IMSS), which is commonly used in the literature (Dougherty and Escobar, 2013; Levy, 2008).[2]

Higher labor tax rates make informality more attractive because informal workers don't have to pay taxes, whereas formal sector workers do. Of course, there are also advantages of formal sector employment (social security benefits for example),

[1] See the Appendix Table 1.E.1 for a complete list of informality definitions for each country.

[2] Illegality, as defined by Levy (2008) refers to the violation of laws regarding social security, firing and severance pay, and labor taxes (which occur jointly). In this paper, we do not focus on this characteristic of informality. Informality in Mexico is also not defined by the size of a firm, the economic sector or activity, neither by the legality of the firm. These may be characteristics that the informal sector share. For example, a large percentage of small firms are not registered with a social security program. Also, the economic sectors are independent of formality and informality. Both can occur in the agriculture, industry and services sector.

but keeping benefits constant, higher labor taxes are thought to increase informality.[3] While this mechanism is fairly straightforward, what we don't know is: (i) what is the size of these effects? (ii) do they apply to all workers equally, or only to certain demographic or income-based groups? These questions are the focus of this paper.

We investigate the effects of labor tax rates on informality using a large rotating panel survey, which allows us to control for individual effects, time effects, lagged income, seasonality and a range of demographic, regional and industry-level factors. Controlling for individual effects is particularly important given preferences regarding informality are likely to be heterogeneous, and are possibly correlated with those affecting incomes and tax rates. We exploit a number of tax reforms in Mexico over 10 years between 2005-14, such as changes in the number and position of tax brackets, and changes in the availability of tax credits. In some years reforms raised average tax rates for low income earners relative to high income earners, and in other years the opposite. Our time effects remove any economy-wide movements, and so we rely on tax changes of some income earners relative to others. As average tax rates depend on income, which is potentially endogenous to informal status, we follow Auten and Carroll (1999) and instrument actual average tax rates with the change in tax rates based on lagged income. To calculate average tax rates for each worker, we manually apply tax rules, as described by the OECD. We calculate hypothetical tax rates for informal workers (based on their reported earnings) and "invert" the schedule to calculate gross income implied by the reported after-tax income in our rotating panel survey.

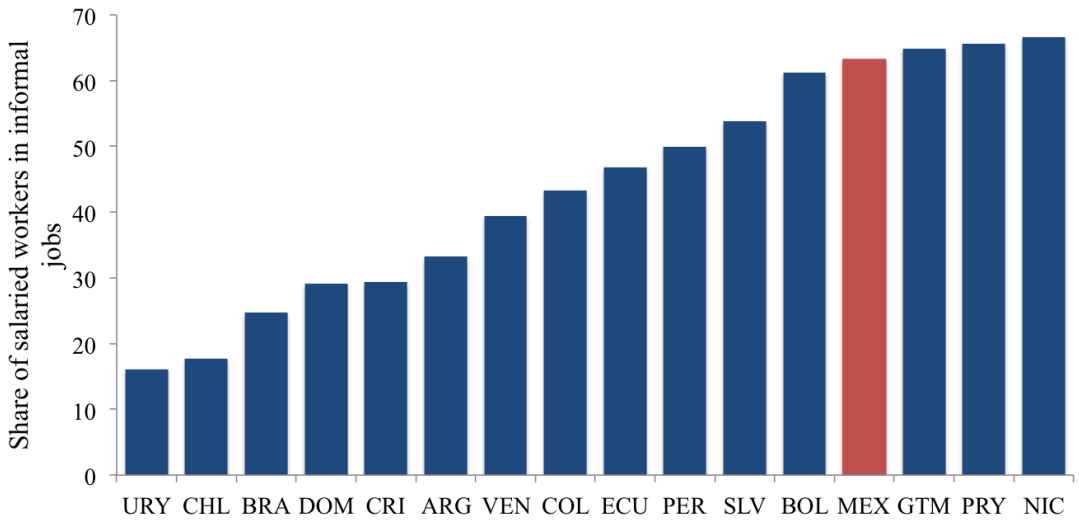
Given our rich dataset, we are able to uncover some new stylized facts about informality in Mexico. By constructing a proxy of the marginal product of labor (MPL) for each worker we can measure how productivity changes when individuals

[3] As informality is an extensive margin decision rather than an intensive margin one, it is important to look at the effect of average tax rates (rather than marginal tax rates). It is also important to include all labor taxes: those on the employer and the employee.

transition from the informal sector to the formal sector (or vice versa). Surprisingly, we find the MPL only increased by about 13 percent (of the average) when workers transition to the formal sector, and falls by a similar amount when they transit out of it. This suggests relatively small static costs of informality, of only around 6.5 percent of GDP. However we do find that the MPL tends to grow around 2 percent faster for formal sector workers than informal workers (even after controlling for education, age, industry and other variables), which suggests that some of the costs are dynamic. Like others in the literature, we find high rates of transition in and out of formality — about 15 percent in and out of informality each year for Mexico. This suggests that there is little segmentation between formal and informal sectors. We also show that although informal sector wages are lower (on average) than formal sector ones, there are many informal workers who earn more than formal sector ones — suggesting that informal work is not just the reserve of low wage occupations or those who were unable to find a job in the formal sector.

Returning to our main question, we find that an increase in average tax rates only results in a statistically significant increase in the probability of being employed in the informal sector for women and low income workers, but not for the population as a whole. This result is consistent with anecdotal evidence that women and people with low income may be closer to being indifferent between working in the formal or informal sector (Chant, 1991; Alter Chen, 2001; Bernal Salazar, 2009). On average a one percentage point increase in the income tax rate implies an increase of around 0.8 percentage points in the probability of becoming informal (or not leaving formality) for women and a 1 percentage point increase for people with low income. The results are strongest for low income women, where a 1 percentage point increase in tax rates increases the probability of a transition to informality by 1.55 percentage points. Further robustness checks also validate these results.

Figure 1.1: **Informal employment is high in most Latin American countries**



Source:CEDLAC and Gasparini and Tornarolli (2009) *Note*: The definition of informality for this figure is the legalistic or social protection definition: A salaried worker is informal if s(he) does not have the right to a pension linked to employment when retired.

Relation to literature To our knowledge, there are very few papers which estimate the effect of taxes on informality. For example, while Anton, Hernandez, and Levy (2013) and Levy (2008) argue that taxes on the formal sector are responsible for encouraging informal employment, they don't estimate the magnitude of these effects.

The closest paper to ours is Schramm (2014), who estimates the equilibrium effects of taxation on sectoral choice, work hours and wages in Mexico, finding that workers are sensitive to income taxation. She shows that around 25 percent of sectoral mobility from the period 1988-2004 is explained by changes in the average tax rate.

Although we look at related issues in a similar context we have very different approaches. Methodologically, Schramm (2014) takes a more structural approach by estimating various elasticities in a model of informality. For example, she calculates participation elasticities, the elasticity of substitution across skill levels, etc. in a large system of equations — many of them estimated at higher levels of aggregation. The effect of tax rates depends on all these elasticities, and may

be model-dependent. In contrast, in this paper we estimate a reduced form relationship between average tax rates and informality that does not depend on the structure of a particular model.

Crucially our samples are very different because Schramm's sample includes only men, whereas our most significant results are for female workers. We also use a different sample of tax changes with Schramm (2014) using a 1988-2004 and our sample running from 2005-2014. Her dataset (the ENEU) only surveys urban workers where as our strongest results are for rural workers (using a different dataset, the ENOE).

The structure of this paper is as follows: Section 2 provides a broader literature review on informality, Section 3 outlines methodology that will help with the understanding on how do the changes in taxes affect the level of informality in Mexico; Section 4 discusses the data used, the tax changes in Mexico and presents three stylized facts on informality in Mexico; Section 5 presents results, and Section 6 concludes.

1.2 Literature review

Background on informality. This section includes the main debates surrounding informality. There are two main theories about informality in an economy. On one side of the spectrum, Fields (1975) and Mazumdar (1976) propose segmentation models, where the supply of formal jobs is scarce, workers are waiting for formal jobs but cannot afford being unemployed, ending up with an informal job. On the other side, Maloney (Maloney, 1999; Maloney, 2004) and Levy (2008) suggest two sector sorting models where informal workers self-select into informal activities because they may earn more in informal self-employment compared to what they could earn in formal jobs; they may also get more independence from self-employment, more flexibility, and may not value protections such as health

insurance which formal work offers them. Alternatively labor conditions have low enforcement of labor regulations.

According to Fields (2009) there is a combination of the two main theories, with a mix of segmented and self-selected individuals among informal workers. Over the past few years, a growing theoretical literature explores models that analyze the effect of different policies on the share of formal employment in the economy. Albrecht, Navarro, and Vroman (2009) claim that the size of the informal sector depends mostly on the worker's productivity levels. By using a model with heterogeneous workers, low-productivity workers decide to become informally self-employed. Mariano. Bosch et al. (2013) show that the effects of policies such as introducing unemployment benefits on informality are larger compared to other analysis where they do calibration exercises. Other models with heterogeneous workers argue that worker's education is one of the main determinants (Korlm and Larsen, 2004), or productivity differences are the most important determinant for changes in informality (Boeri and Garibaldi, 2006). To sum up, the main determinants that this literature includes are the mass of workers and firms that are at the margin between formality and informality; and, how incentives such as policy programs and tax rates change for firms and workers to operate formally (Mariano Bosch and Campos-Vazquez, 2014).

Background on informality in Mexico. Many academics have studied the informal sector in Mexico (Maloney, 2004; Arturo Anton Sarabia, 2013; Jorge Alonso Ortiz, 2011; Anton, Hernandez, and Levy, 2013; Levy, 2008). Mariano Bosch and Campos-Vazquez (2014) analyze the case of a social security program in Mexico for informal workers; Levy (2008), Anton, Hernandez, and Levy (2013), and Matias Busso (2012) estimate that controlling for size and sector of activity at the six-digit level, Mexican informal firms are considerably less productive than formal ones.

Fiess, Fugazza, and Maloney (2010) analyze the cyclical behavior of informality with a two sector model for a number of Latin American countries including Mexico. The results turn out to be more complex than expected, where they confirm that there are episodes where the expansion of informal self-employment is consistent with the traditional segmentation views of informality, but they also identify episodes where informal self-employment behaves "pro-cyclically", driven by relative demand or productivity shocks to the non-tradable sector. Fernandez and Meza (2014) look also at the business cycle in Mexico comparing it with the labor market in Canada. They conclude that the informal market is strongly counter-cyclical.

Ordoñez (2014) analyzes the distortions associated with the presence of an informal sector and incomplete tax enforcement. Using a dynamic general equilibrium framework, he calibrates the model using Mexican data and finds that, under complete enforcement, Mexico's labor productivity and output would be 19 percent higher under perfect competition, and 34 percent higher under monopolistic competition. However, the distortions in Mexico lead to the misallocation of resources towards small and unproductive plants as they engage in tax evasion; distortion in occupational choices, as unproductive entrepreneurs are attracted to the market; and distortions in the capital use of informal establishments, as they reduce their scale to remain undetected.

Dougherty and Escobar (2013) conduct regression analysis to understand the differences in the determinants of informality across Mexico's states. They find that there are multiple factors that explain the differences in informal employment across Mexican states, such as per capita income, quality of labor skills, differences in the prevalence of micro firms, cost to start a business, and rule of law among others.

Levy (2008), Anton, Hernandez, and Levy (2013), and Matias Busso (2012) analyze the implications of Mexico's social insurance system in a context of informality. They argue that one of the causes of poor performance of Mexico's tax

system is the structure of incentives of the benefit systems. The formal sector contributes through payroll taxes for the social benefits, whereas the informal sector receives a bundle of benefits without paying any tax. This disparity leads to a more attractive informal sector. Anton, Hernandez, and Levy (2013) further develop a model to study the implications of the social insurance framework and propose a reform to shift taxation for social insurance from labor to consumption, relying largely on a reformed VAT.

While there have been many studies of the informal sector for the Mexican economy, there is limited evidence on the sensitivity of workers to taxation for developing countries.[4]

Informality and taxes in Mexico. This section includes a discussion of papers that have also done empirical research on the effect of taxes on informality in Mexico. Anton, Hernandez, and Levy (2013) develop a model to study the implications of the social insurance[5] architecture in Mexico in a context of informality and imperfect tax enforcement. They argue that the current framework provides workers with erratic and incomplete coverage against risks, fosters evasion and narrows the tax base, de-links contributions from benefits undermining fiscal sustainability, and distorts the labor market lowering real wages and total factor productivity.

While we do not analyze the distortions on total factor productivity, we also find that the current framework fosters evasion and narrows the tax base. Anton, Hernandez, and Levy (2013) show in their stylized facts that the distribution of employment by formality status is characterized by a skewed size distribution of firms, with a large number of mostly small informal firms evading taxes, on one end, and a very small number of large establishments (mostly formal) complying with taxes on the other end.[6] They also find that Mexico's labor market is characterized by large mobility of workers between formal and informal status. Their research

[4] See Piketty and Saez (2012) and Schramm (2014).
[5] Social insurance is included in labor taxes in Mexico.
[6] Anton, Hernandez, and Levy (2013) use 2008 Census data in their section of stylized facts.

looks at changes only between 2007 and 2008, and show very similar results to what we find by looking also at yearly and quarterly variations in informality.

Levy (2008) focuses on how social protection programs are distorting the labor market. More specifically, subsidies to social security can segment the labor market into a formal and an informal sector, even when there are no barriers to worker's mobility, such as binding minimum wages or others. He argues that because social protection programs subsidize evasion of social security, as informality increases the fiscal constraints under which social policy operates get tighter: more workers receive free social benefits, and the tax base erodes as fewer firms and workers pay taxes or contribute to social security.

Finally, Schramm (2014) estimates the equilibrium effects of taxation on sectoral choice, work hours and wages in Mexico, finding that workers are sensitive to income taxation. She shows that around 25 percent of sectoral mobility from the period 1988-2004 is explained by changes in the average tax rate.

Similarly, we examine the role of fiscal changes in Mexico's recent history (2005-2014) and its effect on the size of the informal sector. We exploit the exogenous variation in the impact of tax rate changes across individuals, by looking at how changes in the average labor tax rates affect the presence of an informal economy. The identification strategy considers that changes in taxes are fairly random, since the government does not target any particular sector of the population (formal or informal). To our knowledge, there hasn't been any research on the effects of changes in labor taxes on informality over the period we estimate.

Our focus is solely on changes in tax rates that are driven by federal taxes.[7] We look at the distribution of wages across individuals, and our sample allows us to follow workers over time. We then estimate how changes in taxes are affecting individuals to move from one sector to the other. Our results differ from previous

[7]Many states have implemented a small income tax in recent years of 1-2%. Taxes currently collected on salaried labor by state governments are unrelated to social benefits. Income from revenue-sharing formulas on taxes collected by the federal government and transfers from the federal budget account for more than 90 percent of state's revenues (Levy, 2008).

research since we look at how changes in taxes are affecting at the individual level the decision of moving from the formal to the informal sector. We find that on average, tax changes increase the probability of an individual moving from the formal to the informal sector.

History of the tax system in Mexico. The tax structure has been changing during the last few decades in Mexico, however the country has not been able to significantly increase revenue from taxes. In general, tax revenue in Mexico represented 13.68 percent of total GDP in 1990, which compared to other Latin American countries is rather low. For example in Chile, it represented around 15.51 percent of GDP, and 21.43 for Brazil in 1990. In 2013, Mexico's tax revenue increased only to 16.62, while in Chile it increased to 18.79, 24.45 percent in Argentina, and 26.51 percent in Brazil.[8] Income Tax Revenue has also been very low for Mexico compared to other countries in the region. In 2013, it was only 5.96 percent of GDP, compared to 7.44 percent in Brazil, and 7.25 percent in Chile. In OECD countries, income tax revenue reaches 11.38 percent in Germany, 12.11 percent in the United States, and 10.91 percent in France.

The 90's were characterized by the entrance of Mexico in the OECD and the adoption of the North America Free Trade Agreement (NAFTA) with the US and Canada. The two events required the country to adopt a different economic structure that facilitated international trade and investment. This new structure was characterized by a new development strategy based on economic openness, deregulation, and privatization, together with an alternative tax structure. The extensive tax reform implemented by President Salinas, consisted of reducing income tax rates applicable to corporations in an effort to set a competitive level against trading partners. The maximum personal income tax rate was also lowered from 50 percent to 35 percent (Bernardi, 2008).

[8]See http://data.imf.org/?sk=77413F1D-1525-450A-A23A-47AEED40FE78 for further detail.

The tax system and informality continue to be at the center of public policy debates in Mexico. For example 2014 tax reforms seek to promote the formalization of the workforce via substantially reducing personal social security charges among other changes.[9] This is the last of a range of income tax changes over the past decade which reformed the number of tax brackets: 5 income brackets from 2005 to 2007, to 8 income brackets between 2008 to 2013, and to 11 income brackets in 2014. See Section 1.C for further details on changes to the tax system over 2005-14.

1.3 Methodology

1.3.1 Conceptual framework and threats to identification

Consider the problem of a worker choosing between working in the formal and the informal sector. If s/he works in the formal sector, there are a number of social benefits, as well as greater job security, which will be idiosyncratically valued by the worker. The availability of formal sector work will also vary with time-invariant individual characteristics such as ability. If the worker chooses to be informal, they (and their employer) do not have to pay taxes, though they also lose some social benefits and may have a different level of productivity. We are interested in estimating the effect of tax rates on informality, as avoiding these taxes is hypothesized in the literature as a main advantage of informal sector work, and is directly under control by the government. Crucially, as the formal/informal decision is at the extensive margin rather than at the *intensive margin*, it is the average tax rate which is the relevant concept (rather than the marginal tax rates).

Given this set up, one can imagine a simple reduced form specification for individual j at time t as in Equation 1.1. I_{jt} is a dummy variable equal to 1

[9]The fiscal reform includes tax reductions that will decrease gradually over time until the firms are incorporated in the general regime after ten years. The individual income tax, or ISR, used to have several exemptions and deductions, which were mainly used by people with higher incomes. Included in the latest changes, it was proposed a series of reforms that would increase the individual ISR tax base by increasing contributions of people in the highest income bracket.

if the individual is informal and zero otherwise. $\tau_t(y_{it})$ is the average tax rate, which includes all taxes on labor (employee income taxes, employee social security taxes, employer social security taxes). The tax rate is a function of the individual's income in that year y_{it}, though the tax code that maps incomes into average tax rates applies to everyone but might change over time (and hence only has a time subscript). $X_j t/4$ are a range of demographic, geographical, and sectoral trends which could influence the formal/informal decision. γ_j is an individual effect, θ_t is a time effect, θ_q is a seasonal effect, and e_{jt} is an iid error capturing all other random factors which might affect the individual's formal/informal sector choice.[10] We are interested in estimating β, the effect of a change in average tax rates on the probability of informality.

$$P(I_{jt} = 1) = \beta\tau_t(y_{it}) + \delta X_j t/4 + \gamma_j + \theta_t + \theta_q + e_{jt} \tag{1.1}$$

Problem 1A (individual heterogeneity): Unobserved characteristics γ_j like culture, experience with informality, and preferences might affect both the productivity y_{it} (and hence average tax rates), as well as the individual's choice over informality. That is $E[\tau_t(y_{ilt} \times \gamma_j] \neq 0$

Problem 1B (seasonality): The seasonal demand for informal work (for example during harvest) θ_q might be correlated with seasonal variation in formal sector productivity (e.g. manufacturing in the lead up to Christmas) which determines tax rates. Our purpose is to *isolate* the effect of changes in average tax rates on the probability of informality only, not because of a worker going from formal to informal sector due to seasonal jobs.

To solve Problem 1 we take four-quarter ended difference (i.e. 2010Q1 less 2009Q1), which removes both the seasonal effect and the individual effect.

[10] θ_q could be included in the the time effects θ_t, but we add them separately to divide macroeconomic mechanisms driving θ_t from seasonality driving θ_q.

$$P(I_{jt} = 1) - P(I_{jt-4} = 1) = \beta \Delta_4 \tau_t(y_{jt}) + \Delta_4 \theta_t + \delta X_j + \Delta_4 e_{jt} \tag{1.2}$$

Problem 2 (time-varying macroeconomic effects): Macroeconomic shocks might affect both tax rates (as the government tries to conduct counter-cyclical policy, or balance its budget), and the ability of the workers to get a job in the formal sector, i.e. $E[\beta \Delta \tau_t(z_{jlt}) \times \Delta_4 \theta_t] \neq 0$

To solve this problem, we include time dummies γ_t in the estimated equation 1.2 to control for variation in $\Delta_4 \theta_t$. This also means that the only variation in tax rates is cross-sectional.

Problem 3 (progressive/regressive tax schedules): As these are federal taxes, the only way they vary across individuals is according to their incomes/productivity y. This means that as taxable income rises (which might be negatively related to informality), a worker's average tax rate may also rise even if the tax schedule has not changed. For example, suppose a worker learns a new skill which is attractive to formal sector employers, then s/he will be less likely to be informal and also the average tax rate paid will change (because the tax schedule is progressive/regressive), generating a spurious correlation between tax rates and the probability of informality.

To see this more clearly, the change in tax rates can be decomposed into a component working only through changes in income/MPL, and a component working only through the tax code. That is, the change in average tax rate is represented by:

$$\Delta_4 \tau_t(y_{jt}) \equiv \tau_t(y_{jt}) - \tau_{t-4}(y_{jt-4}) \tag{1.3}$$

By adding and subtracting $\tau_t(y_{jt-4})$, this can be rearranged as:

$$\begin{aligned}\Delta_4\tau_t(y_{jt}) &\equiv \tau_t(y_{jt}) - \tau_t(y_{jt-4}) + \tau_t(y_{jt-4}) - \tau_{t-4}(y_{jt-4}) \\ &= \underbrace{\bar{\tau}_t(\Delta_4 y_{jt})}_{Effect\ through\ \Delta MPL} + \underbrace{\Delta_4\tau_t(\bar{y}_{jt-4})}_{Effect\ through\ \Delta tax\ rates}\end{aligned} \tag{1.4}$$

The first component is the "endogenous" part operating through changes in income (with tax rates constant)

$$\bar{\tau}_t(\Delta_4 y_{jt}) = \tau_t(y_{jt}) - \tau_t(y_{jt-4}) \tag{1.5}$$

The second part is the "exogenous" component due to purely changes in the tax code, keeping income constant at its previous level. Auten and Carroll (1999) call this the "synthetic" change in taxes

$$\Delta_4\tau_t(\bar{y}_{jt-4}) = \tau_t(y_{jt-4}) - \tau_{t-4}(y_{jt-4}) \tag{1.6}$$

To solve this problem we can use the exogenous component of the tax change $\Delta_4\tau_t(\bar{y}_{jt-4})$ as instrument for the change in the tax rate faced by the worker $\Delta_4\tau_t(y_{jt})$. Changes in tax code $\Delta_4\tau_t(\bar{y}_{jlt-4})$ are determined entirely at the federal level (based on past incomes), they are uncorrelated with all individual productivity shocks determining informality, i.e. the exclusion restriction $E[\bar{\tau}_{jt}(\Delta_4 y_{jlt}) \times e_{jt}] = 0$ is likely to hold (more on this below). Moreover, $\bar{\tau}_{jt}(\Delta_4 y_{jlt})$ is likely to be a very strong instrument for $\Delta_4[\tau_t(y_{jlt})]$ because of Equation 1.4. See Auten and Carroll (1999) for a further discussion on the endogeneity of the tax rate given changes in income.

Problem 4 (omitted variable bias): It is possible that time-varying regional, industry, demographic or income trends X_j affecting informality could be correlated with changes in tax rates in small samples. To get around this problem, we include a range of controls in the specification, described further below.

1.3.2 Empirical strategy and specification

Our empirical strategy follows a difference-in-differences estimator by assessing whether changes in average tax rates affect the probability of informality in the overall population, as well as in more susceptible groups such as women, or poor workers. As argued above, we work in differences to remove individual effects (as well as seasonality), and instrument changes in average tax rates with a "synthetic" tax rate variable only related to the change on the tax code. By construction, the instrument eliminates the effect of income changes attributable to tax-induced behavioral responses on the change in the tax price and only reflects the exogenous statutory change in tax rates. It is this exogenous change in tax rates that is the primary source of identification of the average tax rate in our model. We include time dummies which remove all economy-wide changes in tax rates.

Our difference-in-difference study varies a little from the one typically used in an analysis of policy variables. First, in a standard diff-in-diff study, the policy only changes once (i.e. there is a comparison between pre-treatment and post-treatment). In contrast, here statutory average tax rates change every year which provides additional time series variation. Second, in the standard diff-in-diff study there is a pure treated group (e.g. people eligible for a particular program) and a pure control (people not eligible for the program). Here there is *heterogeneous treatment strength*, because a change in tax rules will affect some tax payers more than others (based on their incomes). Those with smaller changes in average tax rates represent the control group, whereas those with larger changes represent the treatment group.

As with any diff-in-diff study, the key identifying assumption is that, conditioning on observables, the policy change is randomly allocated. While this is difficult to guarantee without a randomized control trial, we see little systematic variation in average tax rates. Figures 1.3-1.5 (in the data section 1.4.2) plot the change in average tax rates vs income for each year. In some years (like 2009-10) high income

earners faced tax hikes and in other years (like 2006-07) they faced tax cuts. In some years there were large discontinuities in tax changes, such as in 2007-08 someone earning around MXN$250,000 pesos would get a 2 percent tax cut, whereas someone earning a few pesos more would get nothing. Similar changes occur at around MXN$130,000 pesos in the same year, and right through 2005-06 when tax brackets were extensively changed. The more contentious tax changes were the increases in tax rates in 2009-2014 that focused on low income earners. This is the result of an increase in the minimum wage (which indexes many allowances and cutoffs in the tax system). Even in this case (i) there are opposite changes in other years, such as a large fall in tax rates for low income earners in 2005-06, (ii) we control for log of lagged income of the worker, and (iii) the discontinuities in the way minimum wages affect tax rates, in particular around MXN$80,000 pesos. Around half of the years have changes in average tax rates focused on low-income earners as the minimum wage changed (which indexes many allowances and cutoffs in the tax system)

Before turning to the empirical specification, we first need to deal with the issue of probability of transition. We assume a *linear probability model* (LPM), where changes in explanatory variables have a linear effect on the change in probability. Econometrically, if is difficult (or impossible) to estimate using alternatives like fixed effects Probit using standard methodologies (instrumental variables also make things more challenging). To estimate a linear probability model, we simply replace the $P(I = 1)$ in Equation 1.2, with a dummy variable for informality. After taking differences, this means that the change in informality variable $\Delta_4 I_{jt}$ can take three values: zero for no change in status; 1 for a transition from formality to informality, or -1 for a transition from informality to formality.

Our specification can be represented in two stages. First we regress the change in the actual average tax rate $\Delta_4 \tau_t(y_{jt})$ on the change in the "synthetic tax price" $\Delta_4 \tau_t(\bar{y}_{jt-4})$ and the other exogenous variables (Equation 1.1). Consistent estimates

of β are obtained by regressing the change in informality on the fitted average tax rate from the first stage, as well as the exogenous controls.[11]

Although the specification is represented in two stages, the estimation of both steps is always estimated together to make sure the standard errors are correct. We also cluster the standard errors at the state level to control for potential serial correlation over the regions.

$$\Delta_4 \tau_t(y_{jt}) = \alpha \Delta_4 \tau_t(\bar{y}_{jt-4}) + \upsilon X_j + \gamma_t \tag{1.7}$$

$$\Delta I_{jt} = \beta \widehat{\Delta_4 \tau_t(y_{jt})} + \delta X_j + \gamma_t + \Delta_4 e_{jt} \tag{1.8}$$

Controls. Finally, the specification includes a wide range of controls X_j which are designed to solve Problem 4 of omitted variable bias. Most important is $log(y_{jt-4})$. We know that the probability of informality (and transitions into/out of informality) depend on income, and that income can also affect tax rates. So we add a control for log of lagged gross income to make sure that the tax rate is not proxying for the income level/productivity of the worker. Age and age squared are also included to control for life cycle effects, and controls for whether female workers are married.

Another group of variables are connected to the regional or sectoral labor market. For example, in a year that the legislature lowered taxes on low-income earners, the poorest regions of Mexico happened to have an economic expansion, which lowered informal employment. This should not lead to inconsistency, but in finite samples might lead to incorrect inference. One could make a similar case for shocks hitting industries with a large share of informal employees, like retail. To reduce the change of biased estimates, X_j includes dummy variables for over 20 industries

[11] We also include the reduced form estimation of the change in informality on the "synthetic" tax rate $\tau_t(\bar{y}_{jt-4})$. Because the first stage estimates of α are generally close (though slightly below) unity, the OLS estimates are slightly smaller.

and more than 35 regions (cities and rural areas) in Mexico.

1.4 Data and Stylized Facts

1.4.1 Data

We have 2 sources of data to look at the labor market effects of tax policy. First, we use micro-data from the Survey of Occupation and Employment (ENOE), from 2005 to 2014 (10 years). The second is the OECD tax data, which we discuss below.

ENOE. The Survey of Occupation and Employment (ENOE)[12] is the survey that the Mexican government relies on for calculating unemployment statistics and the size of the informal sector. The ENOE has been conducted each quarter since 2005Q1 and covers a random sample of approximately 120,260 households. Each household remains in the survey for five consecutive quarters. We use data for 2005-Q1 to 2014Q2 (34 quarters in total).

This sample allows us to observe individuals in each household with respect to their labor characteristics such as earnings, occupation, industry and benefits. As in many other labor surveys, Mexico's ENOE[13] includes a question addressing whether or not the surveyed employee is affiliated with the Mexican Social Security Institute (IMSS), allowing us to classify the workers as formal if the employee is registered with the IMSS and informal otherwise. The survey also includes demographic characteristics for each individual such as metropolitan area, state, gender, age, marital status, and education level.

The use of panel data allows us to control for unobserved heterogeneity which

[12] http://www.inegi.org.mx/est/contenidos/proyectos/encuestas/hogares/regulares/enoe/default.aspx

[13] While there are other surveys that can capture informality in Mexico, such as National Household Income and Expenditure Survey (Encuesta Nacional de Ingresos y Gastos de Hogares, ENIGH), which includes more household characteristics, they lack the panel dimension that allows us to study worker flows, seasonality, and the business cycles.

might be correlated with changes in tax rates. It also allows us to remove seasonality. As discussed below, we calculate the change in informal sector status over the five quarters the individual is in the sample. Because these are consecutive, the first and last quarter are in the same time of the calendar year, which removed seasonal transitions in informality — for example due to harvest or planting. Working in changes rather than levels differences out the individual effect. As we only calculate the annual change for the last observation for each individual, the sample size is only 20 percent of its initial size. As in many other surveys, income data is missing for around 21 percent of surveyed individuals, which substantially reduces our sample size. We also condition on the individual working which leaves around 300,425 household members of legal working age (between 16 to 65 years old) with employment data.

OECD Tax Data. The second source of data is on historical tax rates, collected for the Mexican tax code for 2005-2014 by the OECD.[14] The analysis focuses only on labor income,[15] since it constitutes the largest share of total income for most workers.[16]

Employers withhold provisional tax payments on wage earnings, which they remit to the tax authorities. Employed individuals are also required to make social security contributions, with the amount based on the individual's salary. We use after-tax labor income and the tax schedule to impute pre-tax income ("gross income"), income and payroll tax rates.

Summary Statistics. Table 1.1 includes summary statistics for the sample. For the 10 years of data available, the female and male population have similar characteristics, with an average age of 38 years and around 9-10 years of education (also considering rural areas). Close to 50 percent of the sample is formal, with

[14] http://www.sat.gob.mx/informacion_fiscal/tablas_indicadores/Paginas/default.aspxandOECD

[15] Personal income tax is required of all Mexican residents.

[16] The tax base includes income from wages, pensions, and financial capital, but unfortunately non-labor earnings are not available at the individual or household levels.

about 5 percent lower net income when the sample includes female workers. The average tax rate has been close to 15 percent for that period. Finally, changes in the tax schedule using past income vary between around -5 percent to +2 percent.

Table 1.1: **Summary Statistics**

Variable	Mean	Std. Dev.	Min.	Max.	Obs.
Panel A: Individual Level (Female and Male workers age 16 to 65)					
Age	38	0.112	17	65	299,799
Years completed in school	9.5	4.3	0	23	299,470
Informal	0.43	0.495	0	1	299,799
Net Income	60,750	35,979	7,740	301,350	299,799
Gross Income	63,637	43,200	7,740	302,469	299,799
Average Tax Rate	0.155	0.042	0.082	0.311	299,799
Ch Tax Rate (pastinc)	-0.002	0.011	-0.045	0.017	299,799
Panel B: Individual Level (Male workers age 16 to 65)					
Age	38	0.114	17	65	197,497
Years completed in school	9.12	4.2	0	23	197,290
Informal	0.45	0.50	0	1	197,497
Net Income	63,103	35,940	7,740	300,000	197,497
Gross Income	66,085	43,179	7740	302,469	197,497
Average Tax Rate	0.157	0.042	0.082	0.311	197,497
Ch Tax Rate (pastinc)	-0.002	0.011	-0.045	0.017	197,497
Panel C: Individual Level (Female workers age 16 to 65)					
Age	38	0.108	17	65	102,302
Years completed in school	10.4	4.4	0	23	102,180
Informal	0.37	0.49	0	1	102,302
Net Income	56,208	35,618	7,740	301,350	102,302
Gross Income	58,910	42,848	7,740	302,469	102,302
Average Tax Rate	0.151	0.042	0.082	0.311	102,302
Ch Tax Rate (pastinc)	-0.001	0.011	-0.045	0.017	102,302

Sample trimmed by top and bottom 1% of net income. The sample is restricted to one observation per individual.

1.4.2 Tax Changes in Mexico

The income tax thresholds and rates change frequently during the analyzed period, 2005-2014. Personal income tax is applied after discounting the tax allowance, which includes a yearly holiday bonus and an end-of-year bonus. Informal sector workers evade taxation completely. Figure 1.2 shows the average income tax rate, which includes social security contributions from employers and employees, as a

function of the before tax (gross) wage. An interesting feature of the Mexican tax system is large discontinuities in the *average* tax rate, which suggests a large jump in the tax bill of a worker when they earn an extra peso. For example, at around MXN$88,000 pesos the tax rate jumps by about 3 ppts. This reflects the loss of a tax credit of MXN$2,600 pesos for people earning over MX$88,000. Ideally we would like to use discontinuities econometrically (known as "bunching" in the literature), but this is not possible without administrative data on tax records as our measures of gross income and average tax rates for each individual are only approximate.

Figures 1.3-1.5 provide an illustration of the size of tax rate changes for each year, which is our instrument of the change in the average tax rate faced by each individual. We provide a brief overview here, although in Appendix 1.C we include a detailed narrative of the changes each year.[17]

- From 2005-06 there were major redrafting of the income tax brackets. The number of brackets remained the same, but the tax rates changed for workers earning between MXN$44,000 and MXN$100,000.

- In 2007 there was a tax reform that reduced tax rates, especially for higher earners, from 29 percent to 28 percent, resulting in a decline in tax rates, especially for high income earners.

- In 2008 there was a major tax reform that increased the number of brackets from 5 to 8, with the largest tax rate at 28 percent. The new brackets lowered taxes for workers earning less than MXN$392,000 pesos.

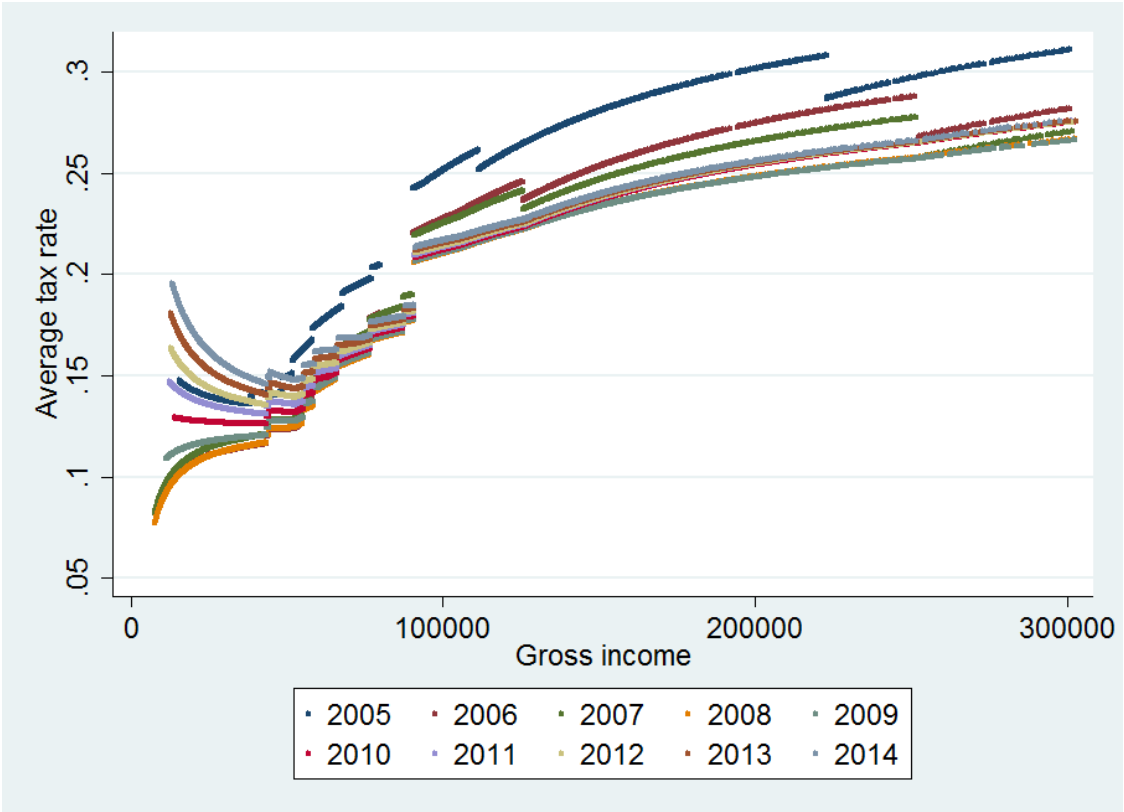
- Figures for 2009, 2011, 2012, 2013 and 2014 all have very similar shapes. In each of these years there were minor changes to statutory tax rates (2009 to 2010 and 2013 to 2014), though the minimum wage did change. Taxpayer's access to a number of tax credits depends on their incomes as multiple of the

[17]The formulas to calculate the average income tax are also explained on the Appendix.

minimum wages, and as the minimum wage increases their average tax rates also change.

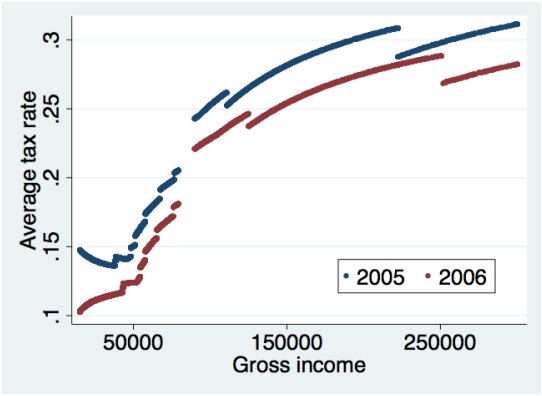
- The figure for 2010 looks like a combination of the figures for 2009 and 2011-14, but with a rate increase for those earning over MXN$123,000. This was a result of an increase in the average tax rate from 19.94 percent to 21.36 percent for workers earning between MXN$123,000 and MXN$249,000. There were also increases in the average tax rate for workers earning between MXN$250,000 and MXN$392,000 pesos.

Figure 1.2: **Overview: Average tax rate (incl. emp SS) & Net income (MXN pesos)**

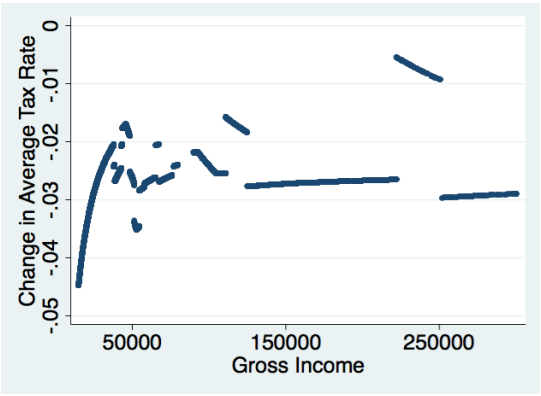


Notes: Gross incomes interpolated for formal sector workers. Data dropped for 363 workers where this algorithm did not converge. See text for further details.

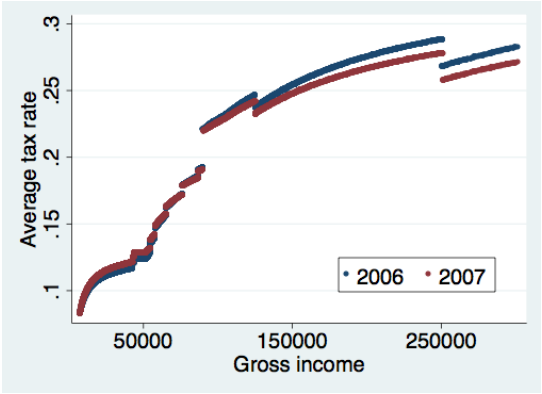
Figure 1.3: **Annual Tax schedule changes (2005-07)**



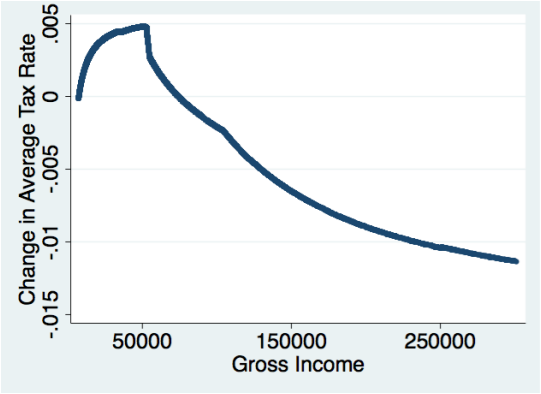
(a) Average tax rates (2005-06)



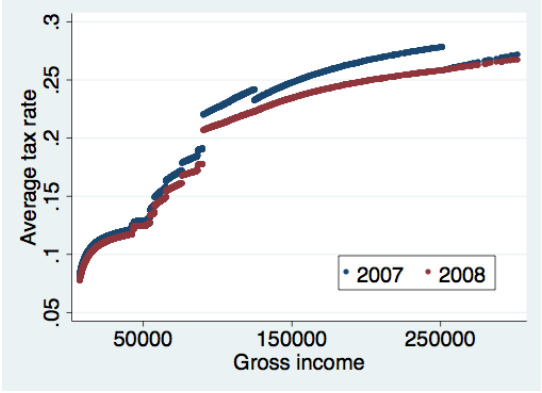
(b) Change in tax rate (2006)



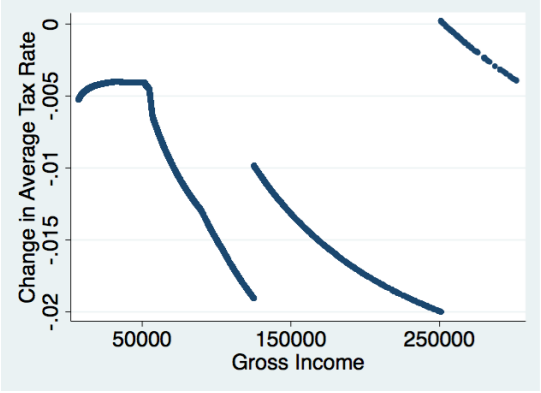
(c) Average tax rates (2006-07)



(d) Change in tax rate (2007)



(e) Average tax rates (2007-08)

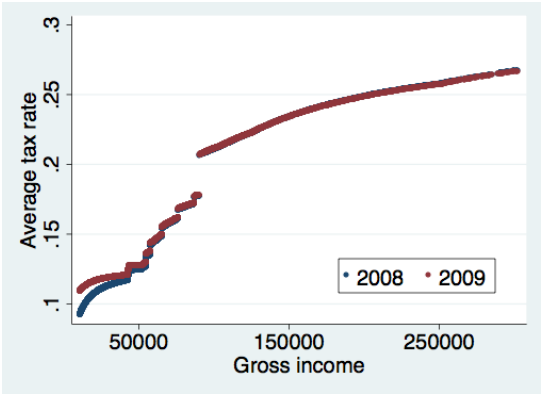


(f) Change in tax rate (2008)

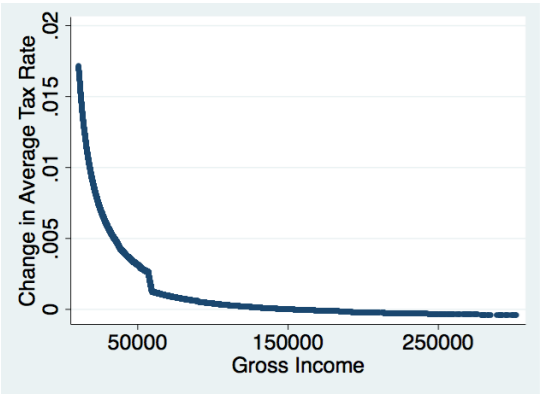
1.4.3 Stylized facts

Formal and informal sector workers are different, but not as much as some might think. Table 1.2, compares a number of demographic characteristics for formal and informal workers and for men and women. The most noticeable differences

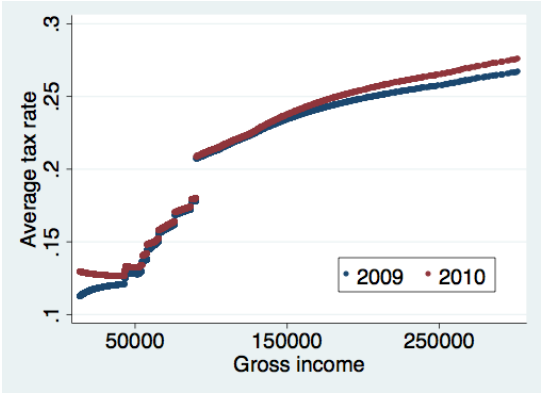
Figure 1.4: **Annual Tax schedule changes (2008-11)**



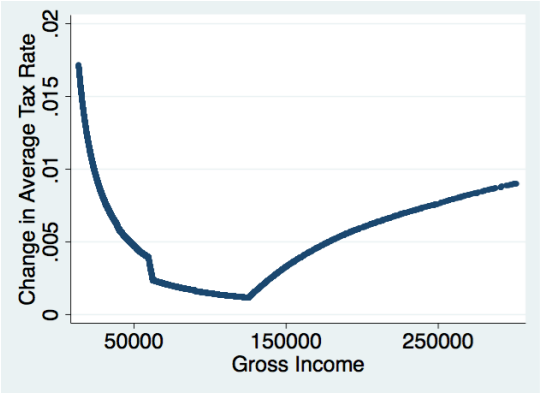
(a) Average tax rates (2008-09)



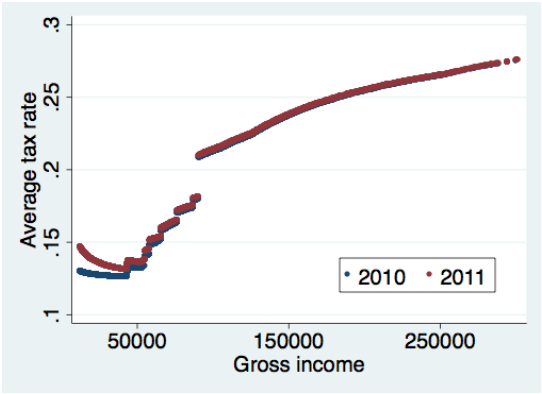
(b) Change in tax rate (2009)



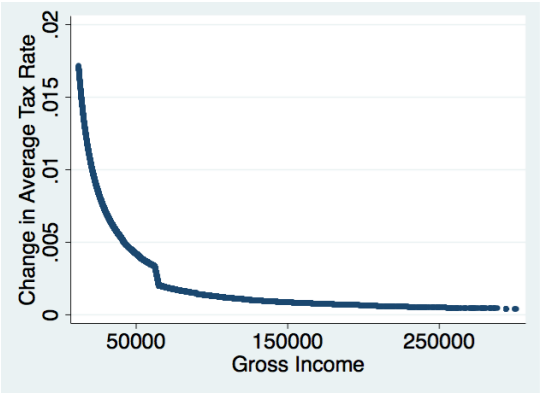
(c) Average tax rates (2009-10)



(d) Change in tax rate (2010)



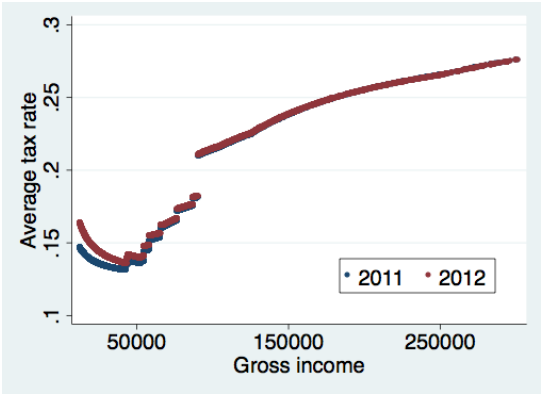
(e) Average tax rates (2010-11)



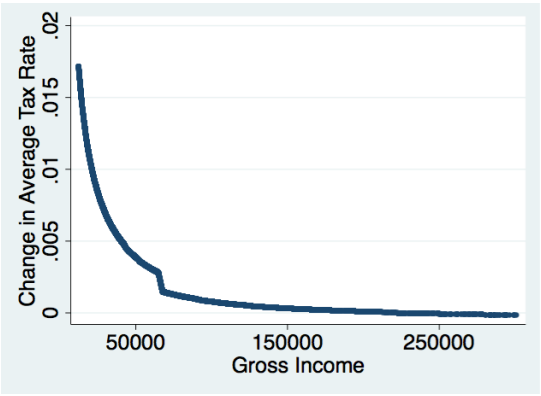
(f) Change in tax rate (2011)

between sectors are the education levels. Women and men working in the informal sector typically have 2-3 years less education than women and men who are in the formal sector. Women in the formal sector are slightly younger than women in the informal sector, while men in the formal sector are slightly older than in the

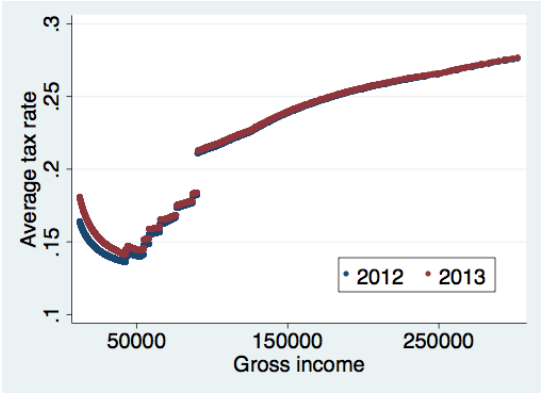
Figure 1.5: **Annual Tax schedule changes (2011-14)**



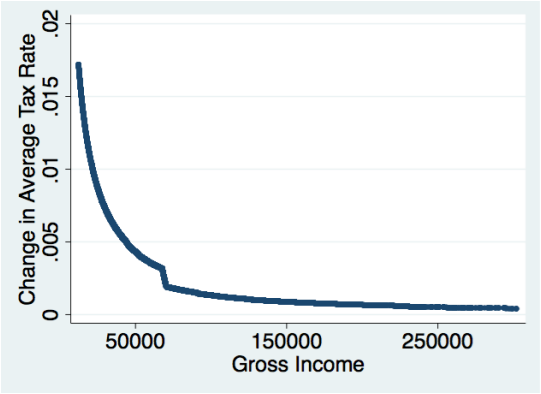
(a) Average tax rates (2011-12)



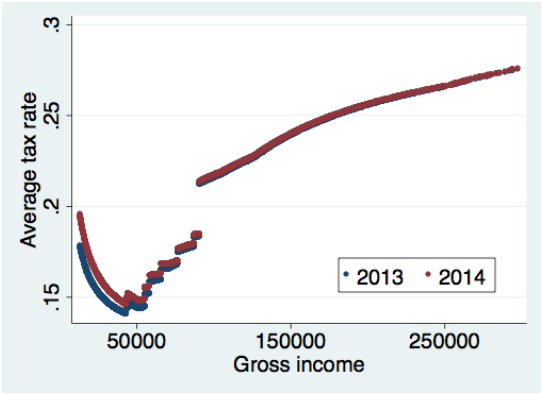
(b) Change in tax rate (2012)



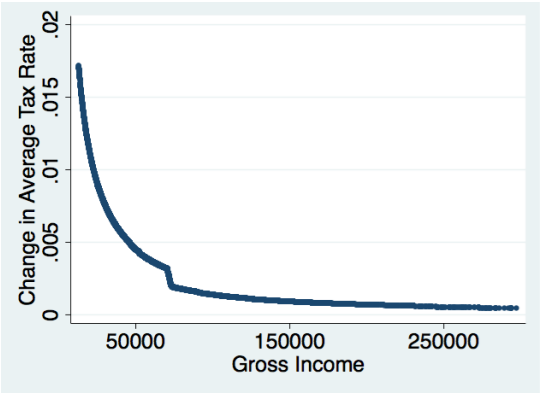
(c) Average tax rates (2012-13)



(d) Change in tax rate (2013)



(e) Average tax rates (2007-2008)



(f) Change in tax rate (2014)

informal sector.

Income and informality

After tax income in the informal sector average MXN$20,000-30,000 less than the formal sector, with a larger gap for women than for men. Pooling across genders, mean income is around MXN$46,000 pesos annually for informal workers, while for formal workers it is around MXN$71,000 pesos. The distribution of formal and informal wages is shown in Figure 1.6. Unsurprisingly the income distribution for informal workers is generally to the left of that for formal sector workers. However, many informal workers earn more than formal workers, which refutes many of the stereotypes that formal sector work is exclusively high paid, and informal sector work is exclusively poorly paid.

Stylized Fact 1: Although average incomes are lower for informal workers than formal sector workers, many informal sector workers earn more than formal sector workers.

Table 1.2: **Sample Statistics by Labor Market State and Gender**

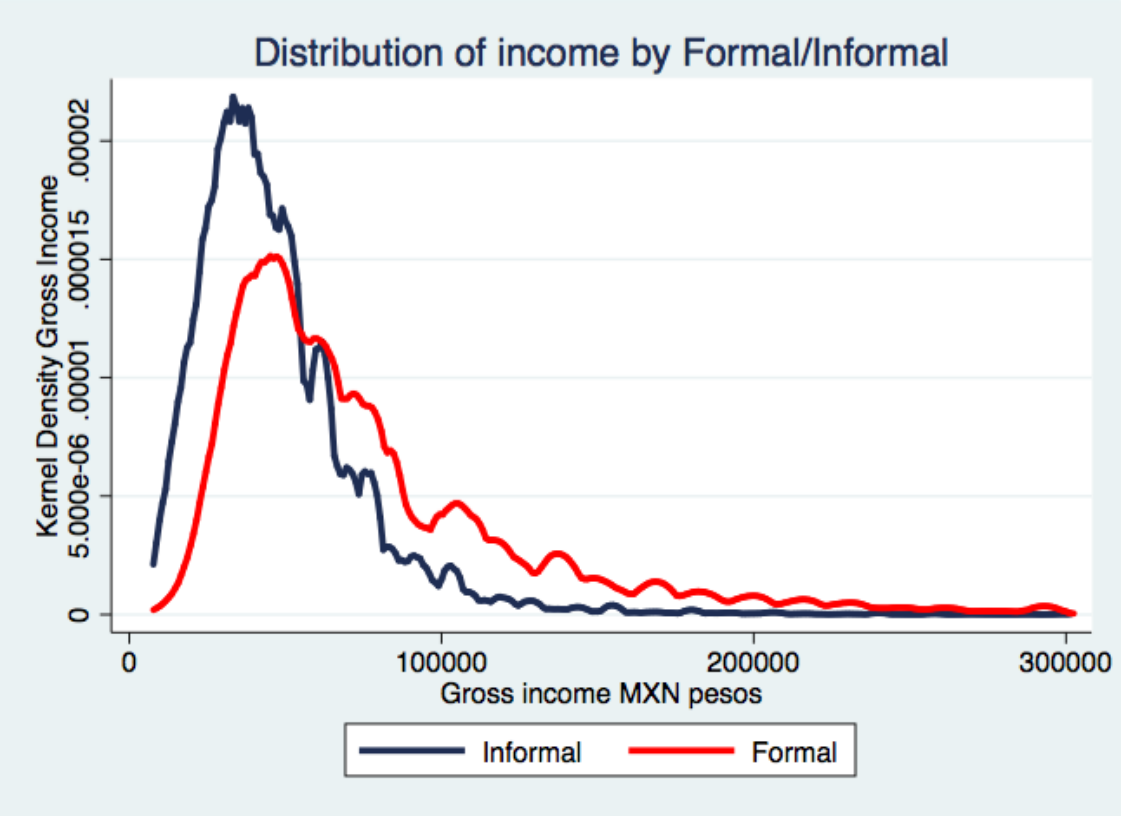
	Men		Women	
Variable	**Informal**	**Formal**	**Informal**	**Formal**
Age	37.8	38.2	38.9	37.8
Years completed in school	7.5	10.4	7.9	11.8
Net Income	51,046	73,075	37,249	67,684
Gross Income	51,046	78,524	37,249	72,021
Average Tax Rate	0.144	0.168	0.132	0.163
Change in Tax Rate (past-income)	-0.001	-0.003	0.001	-0.002
Observations	89,401	108,096	38,574	63,728

Does moving to the formal sector increase wages? The key question for a worker deciding whether to work in the formal or informal sector is their income in each sector.[18]

Conceptually, the fact that wage distributions of formal and informal sector workers are overlapping means that moving from informal to formal sector could

[18] Of course, there are also other considerations, such as the social security and health insurance benefits of working in the formal sector, though as Levy (2008) points out these benefits are weakened now that Mexico has rolled out social benefits for informal workers.

Figure 1.6: **Kernel Density**



lower or raise one's take-home pay. Differences in means are uninformative because, as we have shown in Table 1.2, informal and formal sector workers vary in characteristics like education which determine wages.

To get around this problem we run a regression of the year-ended growth rate of wages of the same worker (e.g. the wage in 2006Q1 relative to 2005Q1) on 3 dummies representing changes in labor market status. The excluded baseline of wage growth for workers always in the formal sector is captured by the constant. By studying changes rather than levels, we "difference out" many of the individual-specific characteristics which might be correlated both with productivity and preference for formal sector work.[19] From a policy perspective, one would

[19] As Levy (2008) discusses, comparing wage rates across workers presents methodological difficulties —most important is the central problem of causal inference since we only observe one of the potential outcomes, the other potential outcome is missing. Second, it is also difficult to identify whether differences in wage rates between formal and informal salaried workers result from differences in labor status or from unobserved variables, even after controlling for age, gender, and years of schooling. Third, in the case of self employed workers, it is difficult to separate earnings into imputed wages and returns to productive assets, and in the case of consultants it is

like to use the change in wage income when informal workers become formal as a partial equilibrium approximation of the economic costs of informality.[20]

Levy (2008) argues that a simple approximation of the increase in output for one worker moving from the informal to formal sectors is $MPL_f - MPL_i = w_f^{AT} + T_f - w_i$ where w_f^{AT} is the after tax formal sector wage, T_f is the taxes paid on formal sector employment and w_i is the informal sector wage. T_f can be divided into three parts: T_f^{IT} (employee income tax); $T_f^{SS_w}$ (employee social security taxes); and $T_f^{SS_e}$ ((employer social security taxes). The gross wage is $w_f = w_f^{AT} + T_f^{IT}$. We construct our measure of $MPL_f = w_f + T_f^{SS_w}$, such that $MPL_f - MPL_i$ is the measure of lost productivity from that reallocated worker. In contrast, Levy (2008) compares w_f^{AT} and w_i, which he rightly argues cannot be interpreted as the productivity differential between formal and informal sector workers.

We construct three measures of the change in the MPL or wage growth. The first is standard change in the gross wage rate for individual j, that is $gw_{j,t} = (w_{j,t} - w_{j,t-4})/w_{j,t-4}$. This is the simplest measure, but does not include the effect of employers payroll taxes, and so is not appropriate for calculating productivity differences. The second measure is similar, but the growth rate of the MPL includes employers social security taxes $gMPL_{j,t} \equiv (MPL_{j,t} - MPL_{j,t-4})/MPL_{t-4}$. The problem is that if $MPL_{j,t-4}$ is "too low" for some reason (such as someone working part time, or because of a reporting error), this will cause the growth rate to "blow up", which has an influence on the estimations. Our preferred measure of the change in the MPL when a worker transfers from the informal to formal sectors is $gaMPL_{j,t} \equiv (MPL_{j,t} - MPL_{j,t-4})/\bar{MPL}_{t-4}$. This measure is very similar to the growth rate in the MPL, but instead of normalizing by last year's *individual* MPL we normalize by the *average* MPL ($\bar{MPL}_{t-4}$), which reduces the sensitivity to outliers.[21] To further reduce the sensitivity to outliers, we drop the top and bottom

difficult to separate imputed wages from compensation for risk or entrepreneurial ability.

[20] Of course there will also be general equilibrium effects as prices change and capital is reallocated, etc.

[21] For example if the household earned MXN$62,500 in 2010Q4, but MXN$6,250 (due to an

5 percent of observations in the default specification, and consider robustness tests dropping the top or bottom 1 percent or 10 percent of observations.

The results using our preferred measure ($gaMPL_{j,t}$) are presented in the final two columns of Table 1.3 and suggest that when workers move from the informal sector to the formal sector, their MPL (as defined above) increases by 13 percent of the average (relative to always informal benchmark).[22]

Likewise, when individuals transfer from formal sector work to informal sector work, their wage falls by around 11-12 percent of the average (relative to always-formal benchmark). Note that the estimated coefficients in the simple specification (Column 5) are very similar to the full specification (Column 6), which includes controls for age, education, number of workers in establishment, or sector (all the observable individual characteristics available from ENOE).

Stylized Fact 2:

1. If a worker moves from the informal to formal sectors, their labor productivity increases by around 13 percent (relative to staying in the informal sector). Likewise if a formal sector worker becomes informal, their MPL falls by 11-12 percent (relative to the always formal worker).

2. Many of the productivity gains of formal sector employment are dynamic. An always-formal worker tends to increase their labor productivity about 2-3 percent faster than an always-informal worker (though this depends on they way the transition is measured).

Levy (2008) argues that the static loss of GDP from informal labor is $(MPL_f - MPL_i) \times \Delta L$, where ΔL is the proportion of workers that would move from informal to formal sector if informal sector work lost its implicit subsidy. An upper bound

reporting error) in 2009Q4, then then increase would be a 900 percent increase in $gMPL_j$, whereas only a 90 percent increase in $gaMPL_j$.

[22] This is calculated as the difference in coefficients between the informal-formal $gaMPL$ and the informal-informal $gaMPL$.

for ΔL is around 50 percent if all informal workers become formal.[23] As such, the *static loss of GDP due to informal employment is at most 6.5 percent GDP*. Of course, there may be second round GE effects, but for a simple partial equilibrium approximation this is remarkably small.

The calculation above also ignores the *dynamic* gains from being formal. The estimates in Table 1.3 suggest that formal sector workers tend to increase their productivity by around 2 percent faster than informal sector workers. This might not sound like much, but if a formal and informal sector worker initially had the same productivity after 35 years the always-formal sector worker would be *twice* as productive as the always-informal sector worker. As around 50 percent of workers are informal, back-of-the-envelope calculations suggest that aggregate labor productivity would grow 1 percent faster if Mexico did not have any informal workers. This finding is robust to controls for a range of socio-economic factors (Column 6), though of course there is always a chance that it could be related to unobserved characteristics of always-formal workers.

Robustness. In Table 1.3 we include two other measures: the growth rate of gross wages gw, and the growth rate of MPL $gMPL$. For informal-formal transitions, the estimated increase in labor productivity is very similar at around 14-15 percent (relative to an always informal baseline), while moving from the formal to informal sector drops MPL by about 12 percent (Column 3-4). Results are quire different when we turn to always-informal wage growth, which is now marginally *higher* than always formal wage growth. As mentioned above, this is probably because this measure puts a lot weight on large increases in the MPL, and so should be treated with caution. All the estimated growth rates of gross wages (gw) are much closer to zero, reflecting the absence of employer social security taxes

[23] An important caveat here is that we are assuming that the gain in MPL applies to all of the 50 percent of informal workers. This is actually a conservative assumption, because transitions from informal $\rightarrow$ formal which drive our estimates of the change in MPL are likely to be those with the largest increases in productivity from this transition. The productivity losses of a large increase in informality could be much much larger because there are unobserved always-formal workers who are much more productive in the formal than informal sectors.

(Columns 1-2). Given the stereotypes around the informal sector, it is striking that gross wage rates are essentially constant when workers move between formal and informal sectors.

In additional results (not reported), we re-estimate Table 1.3 removing 1 percent outliers or 10 percent outliers (rather than 5 percent as in Table 1.3). For $gaMPL$ (our preferred measure), the results are very similar with the alternative outlier classification. Removing 10 percent outliers makes the change in wages in/out of informality a couple of percentage points closer to zero; removing 1 percent outliers makes then a couple of percentage points larger. For the always-informal worker, the 2 percent fall in MPL relative the always-formal is similar in the 10 percent outliers sample, though larger (around 5 percent fall in MPL) in the 1 percent outlier sample. As previously mentioned $gMPL$ and gw are much more sensitive to outliers, especially using only the 1 percent outlier cutoff.

How persistent is informality?

As discussed above, there are two views of informality: one is that informal workers are "segmented" from the rest of the labor market by barriers to entry such as a lack of available formal sector jobs, their cultural background or lack of experience; another is that many workers are close to indifferent to being in the formal or informal sectors, especially given the provision of the benefits for the informal (like Seguro Popular). In the "segmented markets" case, we would expect very little movement between sectors, whereas if workers are close to indifferent then small changes in benefits or wages can make them transition in and out of the informal sector — increasing turnover.

Table 1.4 suggests that transitions in and out of informality are quite high, which is inconsistent with the segmented markets view. Specifically around 14 percent of workers move between status from one year to the next.[24] These results

[24]We are only taking into consideration workers between the ages 16 and 65, and are looking at year-ended changes, which remove seasonal transitions

Table 1.3: **Change in wage rates switching between formal to informal sectors**

Variables	(1) Growth Wage	(2) Growth Wage	(3) gMPL	(4) gMPL	(5) gaMPL	(6) gaMPL
Informal-Informal	0.0067***	0.012***	0.0092***	0.016***	-0.035***	-0.022***
	(0.0012)	(0.0020)	(0.0012)	(0.0019)	(0.00096)	(0.0017)
Formal-Informal	-0.019***	-0.015***	-0.12***	-0.12***	-0.12***	-0.11***
	(0.0024)	(0.0027)	(0.0022)	(0.0025)	(0.0019)	(0.0023)
Informal-Formal	0.041***	0.040***	0.16***	0.16***	0.11***	0.11***
	(0.0023)	(0.0024)	(0.0023)	(0.0024)	(0.0020)	(0.0020)
Age		-0.0029***		-0.0029***		0.00099***
		(0.00031)		(0.00031)		(0.00024)
Squared age		0.000028***		0.000029***		-0.000011***
		(3.99e-06)		(3.91e-06)		(3.07e-06)
Junior high		-0.0012		0.0056		0.0099
		(0.031)		(0.030)		(0.023)
High school		-0.0051		0.0021		0.018
		(0.031)		(0.030)		(0.023)
High school+		-0.0057		0.0017		0.047**
		(0.031)		(0.030)		(0.023)
1-10 workers		-0.013***		-0.013***		-0.018***
		(0.0017)		(0.0017)		(0.0015)
11-50 workers		-0.0050***		-0.0041***		-0.0074***
		(0.0015)		(0.0015)		(0.0014)
Written contract		0.016***		0.013***		0.038***
		(0.0025)		(0.0024)		(0.0021)
No written contract		-0.0068***		-0.0087***		-0.0039**
		(0.0019)		(0.0019)		(0.0016)
Agriculture		-0.010***		-0.010***		-0.028***
		(0.0029)		(0.0029)		(0.0021)
Mining		0.0098		0.0070		0.024***
		(0.0068)		(0.0066)		(0.0074)
Electricity		0.0028		0.00038		0.0086
		(0.0066)		(0.0064)		(0.0073)
Manufacturing		-0.0064***		-0.0062***		-0.027***
		(0.0022)		(0.0021)		(0.0018)
Commerce		0.0038*		0.0032		-0.021***
		(0.0023)		(0.0022)		(0.0019)
Transportation		0.0082***		0.0075**		0.0098***
		(0.0031)		(0.0030)		(0.0027)
Services		0.0027		0.0016		-0.010***
		(0.0020)		(0.0019)		(0.0017)
Urban location		-0.0045***		-0.0042**		0.0033***
		(0.0017)		(0.0017)		(0.0013)
Constant	0.060***	0.14***	0.058***	0.13***	0.072***	0.040*
	(0.00072)	(0.032)	(0.00069)	(0.031)	(0.00071)	(0.023)
Observations	280,171	280,171	280,936	280,936	278,685	278,685
R-squared	0.002	0.003	0.033	0.034	0.034	0.043

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: Standard errors clustered at the state level are shown in parentheses. *** is significant at the 1% level; ** is significant at the 5% level; * is significant at the 10% level. These results drop the top and bottom 5% of wage growth

are similar to other evidence for Mexico. Specifically Anton, Hernandez, and Levy (2013), shows that between 87-91 percent of workers kept their formal/informal status between 2007 and 2008, while around 12 percent of formal workers became informal and 9 percent of informal workers became formal. Our results show that pooling results for 10 years, results are similar — albeit with slightly more transitions. Anton, Hernandez, and Levy (2013) further decomposes these transitions into within-job and between-job, and find that around 6-9 ppts of the transitions between the informal and formal sector (i.e. the majority) involve no change of job.[25]

Stylized Fact 3: Formal and informal status are very fluid in Mexico. Around 15 percent of workers transition between formal and informal sectors each year, *excluding* seasonal transitions, with means around 85 percent of workers remain in their current status.

But those transitions are accentuated by looking at workers classified by gender and education. Our results suggest that women have slightly less mobility compared to men from one year to the next. Specifically, 11 percent of female informal workers move to the formal sector each year, and 13 percent of female formal sector workers become informal. In contrast, these two numbers are 14 percent and 16 percent (respectively) for male workers. Literature suggests (Arturo Anton Sarabia, 2013; Levy, 2008; Maloney, 2004) higher levels of education are correlated with increased skills, hence higher productivity. Using a Logit model, Maloney (2004) shows that workers are less likely to leave formal employment for self-employment (or any other informal sector) as education levels increase. Table 1.4 and Table 1.A.1 show that workers with lower skill levels are likely to transition from formal to informal

[25]Using IMSS data, Levy (2008) describes the informal/formal history of workers. He reports that there are around 15 million changes in worker registration per year, with a total stock of formal workers of around 13 million. Levy (2008) also reports that high income workers have an average stay in the formal sector of 77 percent of their working time (over a ten-year period), while low-wage workers only 49 percent were employed in the formal sector for ten continuous years. On average, he finds that over a ten-year period a worker enrolled in IMSS in 1997, spent 67 percent of the time in the formal sector, and 33 percent of the time in a different labor status.

more often. It could be implied that these workers are the ones at the margin, with income levels that do not change much by being in an informal job or a formal one.

Our main analysis is carried out using only the selection of workers who are employed (formally or informally). We are aware that it could represent a problem if unemployment is also a relevant alternative to employment. Table 1.5 displays the transition matrix among 3 stages: formality, informality and unemployment. Overall, the probability to remain formal if the worker was formal in the first period is of 56 percent. The probability to become unemployed is 28 percent, while the probability of becoming informal is lower, with 15 percent. The results are similar for women and men separately. The Table suggests that the transition to unemployment could be another possible outcome. However, when we include all controls and focus on how changes in taxes could affect the outcome of transitioning to unemployment, the results suggest that changes in taxes do not lead workers to unemployment (see Table 1.10 in the Results section).

Table 1.4: **Annual workers transitions between formality and informality**

Change	$Formal_t$	$Informal_t$
Panel A: Workers status change		
$Formal_{t-1}$	85%	15%
$Informal_{t-1}$	13%	87%
Panel B: Male Workers status change		
$Formal_{t-1}$	84%	16%
$Informal_{t-1}$	14%	86%
Panel C: Female Workers status change		
$Formal_{t-1}$	87%	13%
$Informal_{t-1}$	11%	89%

Table 1.5: **Annual workers transitions between formality and unemployment**

Change	$Formal_t$	$Informal_t$	$Unemployed_t$
Panel A: Workers status change			
$Formal_{t-1}$	57%	15%	28%
$Informal_{t-1}$	12%	58%	30%
$Unemployed_{t-1}$	27%	48%	25%
Panel B: Male Workers status change			
$Formal_{t-1}$	55%	16%	29%
$Informal_{t-1}$	11%	59%	30%
$Unemployed_{t-1}$	26%	50%	24%
Panel C: Female Workers status change			
$Formal_{t-1}$	61%	13%	26%
$Informal_{t-1}$	15%	57%	28%
$Unemployed_{t-1}$	33%	41%	26%

1.5 Results

1.5.1 Main results

In this section, we use regression analysis to assess the effect of changes in the average tax rate on the probability of a worker choosing to work in formal and informal sectors. We first present results for all workers, women and men (Table 1.6), and then break down the sample by income, income and gender, as well as presenting heterogeneous effects with respect to a range of different sub-samples such as by education, age and urban/rural. As mentioned above, we estimate mostly using 2SLS. The first stage is extremely strong: the coefficient on $\tau_{it}(y_{t-4})$ (the instrument) is between 0.7 and 1 with first-stage F-statistics several orders of magnitude above the cutoff of 10.

Table 1.6 shows that for all workers, although there is a significant relation between transitions into informality and average tax rates (Column 1) in simple specifications, it is not robust to controls for time industry and geographic fixed effects (Column 2). Results are similar using a reduced form OLS specification of

informality transitions on tax rates calculated using past income $\tau_{it}(y_{t-4})$.

Gender. The overall insignificant results reflect heterogeneous treatment by gender. For women[26] even when controlling for time, industry and city effects, the results in column (4) and (5) show that a one percentage point increase in the average tax rate also increases the probability of a woman working in the informal labor sector by 0.74 percentage points, significant at the 5 percent level. The OLS results in column (5) are similar, using as the independent variable the "synthetic" change in average tax rates $\tau_{it}(y_{t-4})$. For men however, this result although positive is not significant (columns 6 and 7). In all our results we also control for the lag of gross income (1 year lag) in order to control for the effect of taxes on informality independently of the income level.[27] Our results are consistent with literature findings, which suggest that women are on average more likely to work in the informal sector and are on average more sensitive to an increase in taxes (Gelber and Mitchell, 2012; Eissa and Liebman, 1996; Meghir and Phillips, 2008; Alter Chen, 2001; Bernal Salazar, 2009; Biles, 2009; Chen, 2007).

There have been a large number of studies that have focussed on the effect of taxes on women's labor force participation. An example is Gelber and Mitchell (2012), the authors find that labor force participation among single women increases when "the fraction of their earnings taken away in taxes falls".[28] Similarly, Meghir and Phillips (2008) suggest that taxes and benefits affect differently certain groups of the population. For instance, groups such as women with young children are more sensitive to whether to work or not as well as how many hours they work depending on the levels of taxes and benefits they receive.

A number of other controls are important. The log of lagged gross income is significant and positive. This suggests that workers who previously had higher

[26] For women, we also include a dummy variable for marriage, since the literature suggests that women are more vulnerable to work on the informal sector if there is an increase in taxes (Maloney, 2004; Correia, 2001).

[27] We estimate gross income by using observed after-tax labor income and the tax schedule to impute pre-tax income.

[28] See also Eissa and Liebman (1996) and Meyer and Rosenbaum (2001)

income are more likely to transition to informality, which might be mean reversion after they earned more at a temporary formal job. Higher income workers are generally less likely to be informal, *other things equal*, but this is differenced out when we study changes in informality. Age is positive but age squared is negative, indicating that the probability of being employed informally is inverse-U shaped in age. Being married also seems to have an effect on informal transitions, as do many of industry, time and location dummies included in the model.

Table 1.6: **Main regression results — higher taxes increase informality for women, but not for men or overall**

Informal - Lag(4) Informal							
	Full sample			Women		Men	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	2SLS	2SLS	OLS	2SLS	OLS	2SLS	OLS
Change tax	0.28***	0.15		0.74**		-0.15	
	(0.075)	(0.18)		(0.31)		(0.24)	
Change tax (past income)			0.34*		0.59*		-0.12
			(0.17)		(0.33)		(0.19)
Lag(4) ln Gross Income	0.020***	0.027***	0.026***	0.033***	0.025***	0.024***	0.027***
	(0.0015)	(0.0030)	(0.0016)	(0.0043)	(0.0023)	(0.0044)	(0.0020)
Age		0.44***	0.43***	0.43***	0.47***	0.43***	0.43***
		(0.047)	(0.051)	(0.065)	(0.064)	(0.054)	(0.060)
Age squared		-0.0046***	-0.0044***	-0.0041***	-0.0046***	-0.0047***	-0.0046***
		(0.00059)	(0.00062)	(0.00083)	(0.00076)	(0.00068)	(0.00073)
Married				0.0053***	0.0056***		
				(0.0020)	(0.0020)		
First stage coefficient	0.96***	0.94***		0.88***		0.96***	
	(0.0077)	(0.016)		(0.021)		(0.020)	
Industry FE	NO	YES	YES	YES	YES	YES	YES
Quarter × *Year* FE	NO	YES	YES	YES	YES	YES	YES
City FE	NO	YES	YES	YES	YES	YES	YES
Observations	300,425	303,625	300,425	102,450	112,830	197,975	198,676
F-statistic	15,195.44	3,829.69		1,740.70		2,408.77	

Notes: Standard errors clustered at the state level are shown in parentheses. The OLS results use use the method of Cameron, Gelbach, & Miller (2008) wherein the null distribution of the t-statistic is estimated by running the model in a series of bootstrap samples to deal with a small number of clusters and give more robustness to our results. Variables not shown include industry fixed effects, *quarter* × *year* fixed effects and city fixed effects. *** is significant at the 1% level; ** is significant at the 5% level; * is significant at the 10% level.

Income. From Figure 1.6 we know that informality is concentrated among low-income workers, because it is simply not possible for a number of high-income professional occupations to exist in the shadows.[29]

Table 1.7 shows that average tax rates only have an effect for low income work-

[29]Levy (2008) argues that "capital and technology intensive activities requiring salaried labor are subject to indivisibility and minimum scale, making informal operation prohibitively costly"

ers. When we split the sample at the mean gross income[30] of about MX$62,500 (around USD$4000 exchange rates in early 2015), the results are striking. A one percentage point increase in the average tax increased the probability of being in the informal sector (or reduces the probability of being in the formal sector) by around 1 percentage point for low-income workers (Column 1), but there is no effect for high income workers (Column 2). However, it is not just income in itself which is important, but a combination of income and gender. For lower-income women, a 1ppt increase in average tax rates lead to a 1 to 1.5 percentage point increase in the probability of transition to informality. In contrast, taxes have little effect on transition probabilities for low income men (Column 5). For high-income workers of both genders there is no significant relationship between average tax rates and the transition to informality (Columns 4 and 6).

Table 1.7: **Changes in taxes are contributing to increasing informality specially for the poor**

Informal - Lag(4) Informal						
	Full sample		Women		Men	
	(1)	(2)	(3)	(4)	(5)	(6)
	Income< $MX62{,}400$	Income> $MX62{,}400$	Income< $MX62{,}400$	Income> $MX62{,}400$	Income< $MX62{,}400$	Income> $MX62{,}400$
Change tax	0.95***	0.066	1.47***	-0.65	0.53	0.36
	(0.33)	(0.45)	(0.54)	(0.64)	(0.48)	(0.58)
Lag(4) ln Gross Income	-0.0014	0.031***	0.011***	0.0045	-0.011**	0.043***
	(0.0031)	(0.0099)	(0.0042)	(0.014)	(0.0044)	(0.013)
Married			0.0060**	0.0019		
			(0.0023)	(0.0033)		
Age	0.47***	0.090	0.40***	0.47***	0.53***	-0.064
	(0.047)	(0.12)	(0.072)	(0.17)	(0.056)	(0.14)
Age squared	-0.0049***	-0.00084	-0.0036***	-0.0050**	-0.0058***	0.00087
	(0.00058)	(0.0014)	(0.00091)	(0.0021)	(0.00070)	(0.0017)
First stage coefficient	0.73***	0.88***	0.73***	0.85***	0.72***	0.89***
	(0.016)	(0.046)	(0.018)	(0.074)	(0.025)	(0.047)
Industry FE	YES	YES	YES	YES	YES	YES
$Quarter \times Year$ FE	YES	YES	YES	YES	YES	YES
City FE	YES	YES	YES	YES	YES	YES
Observations	194,529	105,280	71,428	30,876	123,101	74,404
F-statistic	2021.68	355.81	1649.29	132.47	826.35	355.88

Notes: Standard errors clustered at the state level are shown in parentheses. Variables not shown include industry fixed effects, *quarter* × *year* fixed effects and city fixed effects. *** is significant at the 1% level; ** is significant at the 5% level; * is significant at the 10% level.

1.5.2 Robustness tests

So far we have shown that there is some evidence that increases in average income tax rates increase the probability of a transition to the informal sector for women and low-income workers, driven by a strong effect for low-income women. Taxes

[30] We split the sample using gross income of the past period.

seem to have little effect on the informal sector choice of either men or high income workers. In this sections we test the robustness of these findings to other characteristics of workers, such as their age, education level or rural/urban location.

Low-income women. First consider the sample of lower income women, which are the group most sensitive to changes in average tax rates in the results above. In Table 1.8 the first column restates the results from Table 1.7 above (for comparison). In Columns 2 and 3 we see that the results are robust to splitting the sample by years of schooling (with a split at the mean of 10 years). Surprisingly, women with more years of schooling are more sensitive to tax rates, though standard errors are also wider so the results should be interpreted with caution. Columns (4) and (5) split the sample by age at the sample mean of 38 years. Here older women seem to be more sensitive to taxes than younger women. This may also be explained by the literature (Maloney, 2004; Correia, 2001; Chant, 1991; Levy, 2008) which mentions that women may be more flexible, and balance their jobs and families if they work for themselves rather than being employees. Although the coefficient for younger women is insignificant, the point estimates (of unity) is still quite large and so the insignificant coefficient reflects wider standard errors rather than a coefficient of zero. Results are similar for the urban/rural split in Columns 6 and 7. The estimated coefficient of 2.5 (significant at the 1 percent level) is the largest estimated in this paper, suggesting poor women in rural areas are the most sensitive to tax rates — probably because they can transition to informal working in the agricultural sector.[31] As before, the estimated coefficient of unity in urban areas is not small, and insignificance is just driven by wide standard errors.

Women only. Table 1.9 broadens the sample to just including all women. As in low-income women sample, the most sensitive population are for women in rural areas (many of whom are probably of lower income). Many of the other

[31] For rural areas we consider smaller municipalities which are not considered as the main cities in each state. For a complete definition see http://www.inegi.org.mx/est/contenidos/Proyectos/encuestas/hogares/regulares/enoe/doc/Con_basedat2013.pdf

Table 1.8: **Robustness: heterogeneity for sample of low-income women**

Informal - Lag(4) Informal							
	Women with less than MXN 62,400: 2SLS						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
		<10 years in school	>10 years in school	<38 years	>38 years	Urban	Rural
Change tax	1.47***	1.41**	2.24**	0.97	2.10**	0.96	2.46***
	(0.54)	(0.66)	(1.13)	(0.66)	(1.00)	(0.78)	(0.92)
Lag(4) ln Gross Income	0.011***	0.0083*	0.019*	0.019***	0.0063	0.014***	0.0094
	(0.0042)	(0.0043)	(0.011)	(0.0068)	(0.0050)	(0.0049)	(0.0067)
Married	0.0060**	0.0022	0.011***	0.021***	-0.00067	0.0015	0.013**
	(0.0023)	(0.0029)	(0.0044)	(0.0038)	(0.0032)	(0.0038)	(0.0051)
Age	0.40***	0.25**	0.73***			0.43***	0.34**
	(0.072)	(0.10)	(0.18)			(0.088)	(0.14)
Age squared	-0.0036***	-0.0020	-0.0079***			-0.0039***	-0.0031*
	(0.00091)	(0.0012)	(0.0024)			(0.0011)	(0.0018)
First stage coefficient	0.73***	0.69***	0.72***	0.76***	0.67***	0.75***	0.67***
	(0.018)	(0.022)	(0.040)	(0.025)	(0.030)	(0.023)	(0.031)
Industry FE	YES	YES	YES	YES	YES	YES	YES
Quarter × *Year* FE	YES	YES	YES	YES	YES	YES	YES
City FE	YES	YES	YES	YES	YES	YES	YES
Observations	71,428	47,338	24,090	38,741	32,687	45,558	25,870
F-statistic	1649.29	999.73	321.80	897.15	510.47	1057.30	464.39

Notes: Standard errors clustered at the state level are shown in parentheses. Variables not shown include industry fixed effects, *quarter* × *year* fixed effects and city fixed effects. *** is significant at the 1% level; ** is significant at the 5% level; * is significant at the 10% level.

estimates are insignificant (or borderline insignificant) with cuts of the data (by age or education level), which in part reflects larger standard errors but also reflects a small estimated effect for women who are of high income.

We want to note that our results mainly address changes from formal to informal sector jobs. However workers may also leave the labor force when there are increases in taxes. As our results indicate the effect of changes in taxes is more sensitive for women, we also test if changes in taxes increase the probability of leaving the labor force by gender and overall. Table 1.10 shows that the probability of leaving the labor force given an increase in taxes is not significant for either women or men.

Low income only. Table 1.11 cuts the sample by schooling, age and rural/urban for the population of low income workers (both genders). The striking fact is the lack of heterogeneity of point estimates: for example the estimated coefficient is around unity for high and low education workers, and for those in both rural and urban areas — with significance (or lack thereof) being determined by the size of standard errors. The exception is age, where older workers are much more sensitive (in terms of point estimates and significance) to tax rates than younger workers.

Table 1.9: **Robustness: heterogeneity for women only sample**

Informal - Lag(4) Informal							
	Women: 2SLS						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	All women	<10 years in school	>10 years in school	<38 years	>38 years	Urban	Rural
Change tax	0.74**	0.87	0.56	0.54	1.08*	0.37	1.55**
	(0.31)	(0.53)	(0.47)	(0.43)	(0.57)	(0.41)	(0.63)
Lag(4) ln Gross Income	0.033***	0.024***	0.036***	0.038***	0.033***	0.029***	0.039***
	(0.0043)	(0.0066)	(0.0082)	(0.0068)	(0.0071)	(0.0072)	(0.0070)
Married	0.0053***	0.0032	0.0062**	0.019***	-0.00051	0.0018	0.012***
	(0.0020)	(0.0031)	(0.0025)	(0.0029)	(0.0025)	(0.0028)	(0.0041)
Age	0.43***	0.29***	0.71***			0.45***	0.40***
	(0.065)	(0.094)	(0.12)			(0.075)	(0.13)
Age squared	-0.0041***	-0.0023**	-0.0076***			-0.0042***	-0.0038**
	(0.00083)	(0.0011)	(0.0015)			(0.00096)	(0.0016)
First stage coefficient	0.88***	0.87***	0.78***	0.94***	0.81***	0.95***	0.76***
	(0.021)	(0.024)	(0.033)	(0.026)	(0.031)	(0.026)	(0.030)
	(0.021)	(0.024)	(0.033)	(0.027)	(0.032)	(0.026)	(0.032)
Industry FE	YES	YES	YES	YES	YES	YES	YES
$Quarter \times Year$ FE	YES	YES	YES	YES	YES	YES	YES
City FE	YES	YES	YES	YES	YES	YES	YES
Observations	102,304	51,724	50,580	52,656	49,648	67,809	34,495
F-statistic	1790.03	1348.88	545.98	1342.59	665.03	1361.42	621.16

Notes: Standard errors clustered at the state level are shown in parentheses. Variables not shown include industry fixed effects, *quarter* × *year* fixed effects and city fixed effects. *** is significant at the 1% level; ** is significant at the 5% level; * is significant at the 10% level.

Table 1.10: **Robustness: heterogeneity with unemployment as dependent variable**

Unemployment - Lag(4) Unemployment							
	Full sample			Women		Men	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	2SLS	2SLS with controls	OLS	Women 2SLS	Women OLS	Men 2SLS	Men OLS
Change tax	-0.77	-3.93	0.37	-11.4	1.01	-1.73	1.87
	(0.74)	(3.58)	(0.92)	(6.98)	(1.19)	(3.51)	(1.50)
Lag(4) ln Gross Income	0.37***	0.11*	0.18***	0.021	0.18***	0.15**	0.33***
	(0.016)	(0.058)	(0.010)	(0.086)	(0.014)	(0.061)	(0.022)
Married				-0.027			
				(0.027)			
Age		0.96***	0.69***	0.96	-0.78*	0.86***	0.72*
		(0.25)	(0.13)	(0.71)	(0.42)	(0.28)	(0.36)
Age squared		-0.012***	-0.0084***	-0.0097	0.014**	-0.012***	-0.0055
		(0.0035)	(0.0018)	(0.010)	(0.0061)	(0.0038)	(0.0047)
First stage coefficient	0.88***	0.87***	0.78***	0.94***	0.81***	0.95***	0.76***
	(0.021)	(0.024)	(0.033)	(0.026)	(0.031)	(0.026)	(0.030)
Industry FE	YES	YES	YES	YES	YES	YES	YES
$Quarter \times Year$ FE	YES	YES	YES	YES	YES	YES	YES
City FE	YES	YES	YES	YES	YES	YES	YES
Observations	12,145	12,145	13,614	3,466	3,859	8,679	30,664
F-statistic	1172.08	19.87		3.67		25.70	

Notes: Standard errors clustered at the state level are shown in parentheses. The OLS results use use the method of Cameron, Gelbach, & Miller (2008) wherein the null distribution of the t-statistic is estimated by running the model in a series of bootstrap samples to deal with a small number of clusters and give more robustness to our results. Variables not shown include industry fixed effects, *quarter* × *year* fixed effects and city fixed effects. *** is significant at the 1% level; ** is significant at the 5% level; * is significant at the 10% level.

Table 1.11: **Robustness: heterogeneity for sample of all low-income workers**

Informal - Lag(4) Informal							
	Less than MXN 62,400: 2SLS						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	<62,400	<10 years in school	>10 years in school	<38 years	>38 years	Urban	Rural
Change tax	0.95***	0.94**	0.93	0.58	1.30**	1.01*	0.95
	(0.33)	(0.40)	(0.76)	(0.42)	(0.63)	(0.58)	(0.61)
Lag(4) ln Gross Income	-0.0014	-0.0054	0.011	0.0099**	-0.0075*	-0.0035	-0.0021
	(0.0031)	(0.0037)	(0.0075)	(0.0045)	(0.0041)	(0.0040)	(0.0041)
Age	0.47***	0.38***	0.82***			0.56***	0.39***
	(0.047)	(0.047)	(0.12)			(0.081)	(0.069)
Age squared	-0.0049***	-0.0038***	-0.0092***			-0.0058***	-0.0041***
	(0.00058)	(0.00057)	(0.0017)			(0.0010)	(0.00086)
First stage coefficient	0.73***	0.74***	0.68***	0.76***	0.69***	0.75***	0.69***
	(0.016)	(0.018)	(0.037)	(0.018)	(0.024)	(0.015)	(0.023)
Industry FE	YES	YES	YES	YES	YES	YES	YES
$Quarter \times Year$ FE	YES	YES	YES	YES	YES	YES	YES
City FE	YES	YES	YES	YES	YES	YES	YES
Observations	194,529	143,684	50,845	109,950	84,579	108,043	86,486
F-statistic	2021.68	1587.16	350.09	1749.63	835.52	2433.64	906.24

Notes: Standard errors clustered at the state level are shown in parentheses. Variables not shown include industry fixed effects, *quarter* × *year* fixed effects and city fixed effects. *** is significant at the 1% level; ** is significant at the 5% level; * is significant at the 10% level.

Men. Finally, Tables 1.12 and 1.13 corroborate that regardless of education, age or urban/rural location, an increase in the average tax rates for men or low-income men does not increase their probability of working in the informal sector.

Table 1.12: **An increase in the average tax rate does not affect the probability of men to work in the informal market**

Informal - Lag(4) Informal							
	Men: 2SLS						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
		<10 years in school	>10 years in school	<38 years	>38 years	Urban	Rural
Change tax	-0.15	-0.30	-0.27	-0.21	-0.27	0.30	-0.67**
	(0.24)	(0.34)	(0.55)	(0.38)	(0.43)	(0.42)	(0.33)
Lag(4) ln Gross Income	0.024***	0.013*	0.031***	0.032***	0.020***	0.036***	0.013***
	(0.0044)	(0.0072)	(0.011)	(0.0069)	(0.0077)	(0.0088)	(0.0050)
Age	0.43***	0.44***	0.46***			0.44***	0.39***
	(0.054)	(0.053)	(0.11)			(0.088)	(0.080)
Age squared	-0.0047***	-0.0047***	-0.0051***			-0.0047***	-0.0043***
	(0.00068)	(0.00065)	(0.0014)			(0.0011)	(0.00098)
First stage coefficient	0.96***	1.02***	0.82***	0.96***	0.94***	0.99***	0.91***
	(0.020)	(0.027)	(0.031)	(0.028)	(0.026)	(0.024)	(0.032)
Industry FE	YES	YES	YES	YES	YES	YES	YES
$Quarter \times Year$ FE	YES	YES	YES	YES	YES	YES	YES
City FE	YES	YES	YES	YES	YES	YES	YES
Observations	197,505	130,990	66,515	105,933	91,572	112,546	84,959
F-statistic	2215.29	1408.02	695.53	1222.85	1299.26	1766.47	782.63

Notes: Standard errors clustered at the state level are shown in parentheses. Variables not shown include industry fixed effects, *quarter* × *year* fixed effects and city fixed effects. *** is significant at the 1% level; ** is significant at the 5% level; * is significant at the 10% level.

Table 1.13: **An increase in the average tax rate does not affect the probability of poor men to work in the informal market**

Informal - Lag(4) Informal							
	Men with less than MXN 62,400: 2SLS						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
		<10 years in school	>10 years in school	<38 years	>38 years	Urban	Rural
Change tax	0.53	0.72	-0.66	0.18	0.80	0.99	0.21
	(0.48)	(0.58)	(1.30)	(0.63)	(0.96)	(0.94)	(0.70)
Lag(4) ln Gross Income	-0.011**	-0.013**	0.000059	0.0034	-0.018***	-0.020***	-0.0054
	(0.0044)	(0.0055)	(0.011)	(0.0059)	(0.0060)	(0.0069)	(0.0051)
Age	0.53***	0.46***	0.91***			0.65***	0.39***
	(0.056)	(0.055)	(0.19)			(0.12)	(0.078)
Age squared	-0.0058***	-0.0050***	-0.011***			-0.0072***	-0.0043***
	(0.00070)	(0.00068)	(0.0024)			(0.0015)	(0.00098)
First stage coefficient	0.72***	0.74***	0.65***	0.74***	0.68***	0.73***	0.69***
	(0.025)	(0.027)	(0.059)	(0.024)	(0.033)	(0.029)	(0.032)
Industry FE	YES	YES	YES	YES	YES	YES	YES
$Quarter \times Year$ FE	YES	YES	YES	YES	YES	YES	YES
City FE	YES	YES	YES	YES	YES	YES	YES
Observations	123,101	96,346	26,755	71,209	51,892	62,485	60,616
F-statistic	826.35	746.64	123.75	920.78	416.87	646.54	453.88

Notes: Standard errors clustered at the state level are shown in parentheses. Variables not shown include industry fixed effects, $quarter \times year$ fixed effects and city fixed effects. *** is significant at the 1% level; ** is significant at the 5% level; * is significant at the 10% level.

1.6 Conclusions

This chapter has examined how changes in the average tax rate in Mexico affect the decision of an individual to work in the informal vs formal sectors. To answer this question, we use a rotating panel of Mexican workers collected over a decade, calculate implied average tax rates based on the tax code for each worker, and run a regression of individual changes in informality on changes in the average tax rate. Our methodology relies on changes in legislative tax rates as instrument for changes in actual average tax rates, and controls for lagged income, individual effects, time effects, seasonality and a range of demographic, regional and industry-level factors.

Although we find little overall evidence that an increase in the average tax rate affects transitions to informality overall, we do find that an increase in average tax rates induces women and low-income workers towards the informal sector. The results are driven by low-income female workers, with particularly strong results for those in rural areas. Quantitatively, we find that a 1 percentage point in average labor tax rates increases the probability that low income women will transition to informal sector by around 1.5 percentage points. These results suggest that low

income women might be closest to being indifferent between informal and formal sector work.

More broadly, the determinants of informality are important because they might explain why countries like Mexico have experienced such weak productivity growth (Levy, 2008). Our data suggests that although the static costs of informality are smaller than expected, the dynamic costs of informality might be much larger as informal sector workers have slower growth in productivity than formal sector workers. As results suggest that low-income women are particularly sensitive to tax rates, there may be scope for measures in the tax system to encourage these workers back towards formal sector employment.

Appendices

1.A Quarterly transitions between formality and informality

Table 1.A.1: **Quarterly transitions between formality and informality**

Change	$Formal_t$	$Informal_t$
Panel A: Workers status change for full sample		
$Formal_{t-1}$	86.7%	13.3%
$Informal_{t-1}$	11.7%	88.4%
Panel B: Male Workers status change		
$Formal_{t-1}$	86%	14%
$Informal_{t-1}$	12.5%	87.5%
Panel C: Female Workers status change		
$Formal_{t-1}$	88.2%	11.8%
$Informal_{t-1}$	9.9%	90.1%
Panel D: Workers status change with 12 years or less of schooling		
$Formal_{t-1}$	87.9%	12.1%
$Informal_{t-1}$	13.2%	86.8%
Panel E: Workers status change with 9 years or less of schooling		
$Formal_{t-1}$	89.3%	10.8%
$Informal_{t-1}$	15.4%	84.6%

1.B Formulas to calculate average tax rates

Tax allowance:

$$taxal = min(grossinc, min(grossinc * (6/365) * .25, minwage * 15) + min(grossinc * 15/365), minwage * 30)) \quad (1.9)$$

Tax income:

$$taxcr = taxcredit + adj_{fac} * (other_{credit} + extra_{credit} * tax) * (taxinc - llimit)$$
$$if year = 2006 - 2007 \quad (1.10)$$

$$taxcr = taxcredit$$
$$if year = 2008 - 2014 \quad (1.11)$$

Tax paid to government (subtracting credits):

$$taxpaid = taxinc - taxcr \quad (1.12)$$

Employees' social security contributions:

$$ssc_{2005} = grossinc * sscrate + (abs(grossinc - (3 * 365 * minwage))) * sscrate_{sur}$$
$$if year = 2005 \quad (1.13)$$

$$ssc_{2006} = min(grossinc * sscrate, minwage * 25 * 365 * sscrate) + min(abs(grossinc - (3 * 365 * minwage)) * sscrate_{sur}, minwage * (25 - 3) * 365 * sscrate_{sur}) + min(grossinc * sscrate_{dli}, minwage * 23 * 365 * sscrate_{dli}),$$
$$if year = 2006 \quad (1.14)$$

$$ssc_{2007} = min(grossinc * sscrate, minwage * 25 * 365 * sscrate) + min(abs(grossinc - (3 * 365 * minwage)) * sscrate_{sur}, minwage * (25 - 3) * 365 * sscrate_{sur}) + min(grossinc * sscrate_{dli}, minwage * 24 * 365 * sscrate_{dli}),$$

$$if year = 2007 \quad (1.15)$$

$$\begin{aligned} ssc = min(grossinc * sscrate, minwage * 25 * 365 * sscrate) \\ + min(abs(grossinc - (3 * 365 * minwage)) * sscrate_{sur}, \\ minwage * (25 - 3) * 365 * sscrate_{sur}), \end{aligned}$$

$$if year = 2008 - 2014 \quad (1.16)$$

Central government income tax paid plus employees' social security contributions:

$$tgpaid = taxpaid + ssc \quad (1.17)$$

bigskip **Take-home pay:**

$$earn_{af} = grossinc - tgpaid \quad (1.18)$$

Employers' social security contributions:

$$\begin{aligned} sscemp_{2005} = grossinc * sscempr \\ +(365*minwage)*sscempr_{min}+(abs(grossinc-(3*365*minwage)))*sscempr_{sur} \end{aligned}$$

$$if year = 2005 \quad (1.19)$$

$$\begin{aligned} sscemp_{2006} = min(grossinc * sscempr, minwage * 25 * 365 * sscrate) \\ + 365 * minwage * sscempr_{min} \\ + min(abs(grossinc - (3 * 365 * minwage)) * sscempr_{sur}, \\ minwage * (25 - 3) * 365 * sscempr_{sur}) \\ + min(grossinc * sscempr_{dli}, minwage * 23 * 365 * sscempr_{dli}), \end{aligned}$$

$$if year = 2006 \quad (1.20)$$

$$sscemp_{2007} = min(grossinc * sscempr, minwage * 25 * 365 * sscrate)$$

$$+ 365 * minwage * sscempr_{min}$$
$$+ min(abs(grossinc - (3 * 365 * minwage)) * sscempr_{sur},$$
$$minwage * (25 - 3) * 365 * sscempr_{sur})$$
$$+ min(grossinc * sscempr_{dli}, minwage * 24 * 365 * sscempr_{dli}),$$
$$if year = 2007 \quad (1.21)$$

$$sscemp = min(grossinc * sscempr, minwage * 25 * 365 * sscempr)$$
$$+ 365 * minwage * sscempr_{min}$$
$$+ min(abs(grossinc - (3 * 365 * minwage)) * sscempr_{sur},$$
$$minwage * (25 - 3) * 365 * sscempr_{sur}),$$
$$if year = 2008 - 2014 \quad (1.22)$$

Average Income Tax:

$$taxrateave = (sscemp + tgpaid)/grossincome \quad (1.23)$$

1.C Narrative discussion of changes in tax rates

The main changes in tax rates over the period studied depend on the increase in the minimum wage every year.

- 2005-2006: Main changes include a decrease in the last tax bracket from 30 percent to 29 percent and the inclusion of a ceiling to the salary that is base for the social security contributions in 2006.
- 2006-2007: Main changes include a decrease in the last tax bracket form 29 percent to 28 percent. The base for social security contributions changed the ceiling from a base of 23 times the minimum wage to a base of 24 times the minimum wage.
- 2007-2008: Main changes include an increase in tax brackets from 5 to 8, and the fiscal subsidy disappeared, keeping only tax credits. The adjustment factor that accounted for the share of fringe benefits disappeared. The base

for social security contributions changed the ceiling from a base of 24 times the minimum wage to a base of 25 times minimum wage.

- 2008-2009: Main change was a decrease of insurance for work injuries of employees from 2.12 percent to 2.04 percent of worker's monthly wage.[32]
- 2009-2010: Main change was a modification on the last 3 brackets of tax rate. From 19.94 to 21.36 percent; 21.95 to 23.525 percent; and from 28 to 30 percent, respectively.
- 2010-2011: No changes
- 2011-2012: The main change was on insurance for work injuries, which went from 2.05 to 1.98 percent of worker's monthly wage.
- 2012-2013: No changes
- 2013-2014: There were 3 new brackets included: 32 percent, 34 percent and 35 percent for high income earners.

[32]The amount of the work injury fee depends on the risk level in which the company is classified. The Mexican Institute of Social Security provided a weighted average rate that considers the economic activities from C to K of the International Standard Classification (Economic Co-operation and Development, 2011)

1.D First stage of Two-stage Least Squares Regressions

Table 1.D.1: **Main regression results – higher taxes increase informality for women, but not for men or overall – First Stage**

Change tax rate				
	(1)	(2)	(3)	(4)
	2SLS	2SLS	2SLS	2SLS
	Full sample	Full sample	Women	Men
Change tax (past income)	0.96***	0.94***	0.88***	0.96***
	(0.0077)	(0.016)	(0.021)	(0.020)
Lag(4) ln Gross Income	-0.0079***	-0.011***	-0.0097***	-0.014***
	(0.00041)	(0.00046)	(0.00043)	(0.00062)
Age		0.052***	0.029***	0.074***
		(0.0029)	(0.0057)	(0.0038)
Age squared		-0.00065***	-0.00039***	-0.00090***
		(0.000035)	(0.000072)	(0.000047)
Married			0.00031**	
			(0.00015)	
Constant	0.087***	0.11***	0.095***	0.13***
	(0.0046)	(0.0046)	(0.0039)	(0.0060)
Industry FE	NO	YES	YES	YES
$Quarter \times Year$ FE	NO	YES	YES	YES
City FE	NO	YES	YES	YES
Observations	300,425	300,425	102,450	197,975
R-squared	0.220	0.237	0.272	0.231

Notes: Standard errors clustered at the state level are shown in parentheses. Variables not shown include industry fixed effects, *quarter* × *year* fixed effects and city fixed effects. *** is significant at the 1% level; ** is significant at the 5% level; * is significant at the 10% level.

Table 1.D.2: **Changes in taxes are contributing to increasing informality specially for the poor – First Stage**

Change tax rate						
	Full sample		Women		Men	
	(1)	(2)	(3)	(4)	(5)	(6)
	Income<62,400	Income>62,400	Income<62,400	Income>62,400	Income<62,400	Income>62,400
Change tax (past income)	0.73***	0.88***	0.73***	0.85***	0.72***	0.89***
	(0.016)	(0.046)	(0.018)	(0.074)	(0.025)	(0.047)
Lag(4) ln Gross Income	0.00075***	-0.021***	0.00059**	-0.021***	-0.00066*	-0.021***
	(0.00021)	(0.00075)	(0.00028)	(0.00096)	(0.00034)	(0.00077)
Married			0.00056***	0.00062		
			(0.00012)	(0.00039)		
Age	0.040***	0.091***	0.021***	0.053**	0.059***	0.10***
	(0.0026)	(0.0077)	(0.0037)	(0.019)	(0.0040)	(0.0080)
Age squared	-0.00048***	-0.0010***	-0.00027***	-0.00057**	-0.00069***	-0.0012***
	(0.000031)	(0.000088)	(0.000045)	(0.00023)	(0.000049)	(0.000091)
Constant	-0.019***	0.20***	-0.016***	0.22***	-0.0078**	0.20***
	(0.0022)	(0.0083)	(0.0030)	(0.011)	(0.0032)	(0.0086)
Industry FE	YES	YES	YES	YES	YES	YES
$Quarter \times Year$ FE	YES	YES	YES	YES	YES	YES
City FE	YES	YES	YES	YES	YES	YES
Observations	194,965	105,460	71,528	30,922	123,437	74,538
R-squared	0.354	0.140	0.425	0.159	0.329	0.136

Notes: Standard errors clustered at the state level are shown in parentheses. Variables not shown include industry fixed effects, *quarter* × *year* fixed effects and city fixed effects. *** is significant at the 1% level; ** is significant at the 5% level; * is significant at the 10% level.

Table 1.D.3: **Robustness: heterogeneity for sample of low-income women – First Stage**

Change tax rate							
	Women with less than MXN 62,400: 2SLS						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
		<10 years in school	>10 years in school	<38 years	>38 years	Urban	Rural
Change tax (past income)	0.73***	0.69***	0.72***	0.76***	0.67***	0.75***	0.67***
	(0.018)	(0.022)	(0.040)	(0.025)	(0.030)	(0.023)	(0.031)
Lag(4) ln Gross Income	0.00059**	-0.0011***	0.0023***	0.00098***	0.00023	0.00073**	0.00023
	(0.00028)	(0.00032)	(0.00039)	(0.00029)	(0.00039)	(0.00032)	(0.00038)
Married	0.00056***	0.00030**	0.00059**	0.0010***	0.00045***	0.00044***	0.00076***
	(0.00012)	(0.00012)	(0.00023)	(0.00017)	(0.00016)	(0.00015)	(0.00017)
Age	0.021***	0.016***	0.033***			0.022***	0.019***
	(0.0037)	(0.0030)	(0.0090)			(0.0050)	(0.0048)
Age squared	-0.00027***	-0.00018***	-0.00035***			-0.00029***	-0.00022***
	(0.000045)	(0.000040)	(0.00011)			(0.000060)	(0.000062)
Constant	-0.016***	0.00085	-0.034***	-0.015***	-0.011***	-0.016***	-0.014***
	(0.0030)	(0.0032)	(0.0039)	(0.0030)	(0.0038)	(0.0035)	(0.0038)
Industry FE	YES	YES	YES	YES	YES	YES	YES
$Quarter \times Year$ FE	YES	YES	YES	YES	YES	YES	YES
City FE	YES	YES	YES	YES	YES	YES	YES
Observations	71,528	47,401	24,127	38,800	32,728	45,624	25,904
R-squared	0.425	0.537	0.295	0.444	0.407	0.410	0.460

Notes: Standard errors clustered at the state level are shown in parentheses. Variables not shown include industry fixed effects, *quarter* × *year* fixed effects and city fixed effects. *** is significant at the 1% level; ** is significant at the 5% level; * is significant at the 10% level.

Table 1.D.4: **Robustness: heterogeneity for women only sample – First Stage**

Change tax rate							
	Women: 2SLS						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	All women	<10 years in school	>10 years in school	<38 years	>38 years	Urban	Rural
Change tax (past income)	0.88***	0.87***	0.78***	0.94***	0.81***	0.95***	0.76***
	(0.021)	(0.024)	(0.033)	(0.027)	(0.032)	(0.026)	(0.032)
Lag(4) ln Gross Income	-0.0097***	-0.0083***	-0.013***	-0.0088***	-0.011***	-0.010***	-0.0084***
	(0.00043)	(0.00039)	(0.00058)	(0.00050)	(0.00040)	(0.00046)	(0.00052)
Married	0.00031**	-0.00012	0.00050**	0.00091***	0.00020	0.00026	0.00038
	(0.00015)	(0.00014)	(0.00023)	(0.00017)	(0.00020)	(0.00018)	(0.00028)
Age	0.029***	0.014***	0.056***			0.030***	0.028***
	(0.0057)	(0.0040)	(0.010)			(0.0070)	(0.0064)
Age squared	-0.00039***	-0.00020***	-0.00061***			-0.00042***	-0.00035***
	(0.000072)	(0.000053)	(0.00013)			(0.000088)	(0.000081)
Constant	0.095***	0.081***	0.13***	0.093***	0.11***	0.11***	0.078***
	(0.0039)	(0.0038)	(0.0055)	(0.0052)	(0.0040)	(0.0045)	(0.0048)
Industry FE	YES	YES	YES	YES	YES	YES	YES
Quarter × *Year* FE	YES	YES	YES	YES	YES	YES	YES
City FE	YES	YES	YES	YES	YES	YES	YES
Observations	102,450	51,794	50,656	52,743	49,707	67,907	34,543
R-squared	0.272	0.424	0.208	0.280	0.266	0.260	0.304

Notes: Standard errors clustered at the state level are shown in parentheses. Variables not shown include industry fixed effects, *quarter* × *year* fixed effects and city fixed effects. *** is significant at the 1% level; ** is significant at the 5% level; * is significant at the 10% level.

Table 1.D.5: **Robustness: heterogeneity for sample of all low-income workers – First Stage**

Change tax rate							
	Less than MXN 62,400: 2SLS						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	<62,400	<10 years in school	>10 years in school	<38 years	>38 years	Urban	Rural
Change tax (past income)	0.73***	0.74***	0.68***	0.76***	0.69***	0.75***	0.69***
	(0.016)	(0.018)	(0.037)	(0.018)	(0.024)	(0.015)	(0.023)
Lag(4) ln Gross Income	0.00075***	-0.00029	0.0025***	0.0011***	0.00086***	0.0011***	0.00034
	(0.00021)	(0.00022)	(0.00042)	(0.00027)	(0.00022)	(0.00030)	(0.00029)
Age	0.040***	0.032***	0.060***			0.050***	0.029***
	(0.0026)	(0.0025)	(0.0051)			(0.0034)	(0.0029)
Age squared	-0.00048***	-0.00036***	-0.00066***			-0.00061***	-0.00031***
	(0.000031)	(0.000031)	(0.000065)			(0.000042)	(0.000035)
Constant	-0.019***	-0.0072***	-0.041***	-0.014***	-0.013***	-0.024***	-0.014***
	(0.0022)	(0.0021)	(0.0044)	(0.0028)	(0.0024)	(0.0031)	(0.0027)
Industry FE	YES	YES	YES	YES	YES	YES	YES
Quarter × *Year* FE	YES	YES	YES	YES	YES	YES	YES
City FE	YES	YES	YES	YES	YES	YES	YES
Observations	194,965	144,015	50,950	110,172	84,793	108,294	86,671
R-squared	0.354	0.406	0.258	0.374	0.329	0.332	0.391

Notes: Standard errors clustered at the state level are shown in parentheses. Variables not shown include industry fixed effects, *quarter* × *year* fixed effects and city fixed effects. *** is significant at the 1% level; ** is significant at the 5% level; * is significant at the 10% level.

Table 1.D.6: **An increase in the average tax rate does not affect the probability of men to work in the informal market – First stage**

Change tax rate							
	Men: 2SLS						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
		<10 years in school	>10 years in school	<38 years	>38 years	Urban	Rural
Change tax (past income)	0.96***	1.02***	0.82***	0.96***	0.94***	0.99***	0.91***
	(0.020)	(0.027)	(0.031)	(0.028)	(0.026)	(0.024)	(0.032)
Lag(4) ln Gross Income	-0.014***	-0.014***	-0.016***	-0.013***	-0.014***	-0.016***	-0.011***
	(0.00062)	(0.00077)	(0.00060)	(0.00066)	(0.00062)	(0.00072)	(0.00061)
Age	0.064***	0.053***	0.087***			0.083***	0.045***
	(0.0036)	(0.0046)	(0.0076)			(0.0044)	(0.0053)
Age squared	-0.00079***	-0.00066***	-0.00097***			-0.0010***	-0.00055***
	(0.000045)	(0.000056)	(0.00010)			(0.000055)	(0.000062)
Married	0.0015***	0.0014***	0.0016***	0.0023***	0.0016***	0.0022***	0.00069***
	(0.00016)	(0.00017)	(0.00029)	(0.00020)	(0.00022)	(0.00021)	(0.00019)
Constant	0.13***	0.14***	0.15***	0.14***	0.14***	0.16***	0.10***
	(0.0062)	(0.0076)	(0.0061)	(0.0070)	(0.0065)	(0.0075)	(0.0060)
Industry FE	YES	YES	YES	YES	YES	YES	YES
$Quarter \times Year$ FE	YES	YES	YES	YES	YES	YES	YES
City FE	YES	YES	YES	YES	YES	YES	YES
Observations	197,975	131,335	66,640	106,157	91,818	112,826	85,149
R-squared	0.231	0.270	0.195	0.241	0.220	0.224	0.252

Notes: Standard errors clustered at the state level are shown in parentheses. Variables not shown include industry fixed effects, *quarter* × *year* fixed effects and city fixed effects. *** is significant at the 1% level; ** is significant at the 5% level; * is significant at the 10% level.

Table 1.D.7: **An increase in the average tax rate does not affect the probability of low income men to work in the informal market – First Stage**

Change tax rate							
	Men with less than MXN 62,400: 2SLS						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
		<10 years in school	>10 years in school	<38 years	>38 years	Urban	Rural
Change tax (past income)	0.72***	0.74***	0.65***	0.74***	0.68***	0.73***	0.69***
	(0.025)	(0.027)	(0.059)	(0.024)	(0.033)	(0.029)	(0.032)
Lag(4) ln Gross Income	-0.00066*	-0.0015***	0.0011*	0.000040	-0.00036	-0.0013***	-0.00032
	(0.00034)	(0.00035)	(0.00063)	(0.00040)	(0.00033)	(0.00048)	(0.00039)
Age	0.059***	0.049***	0.092***			0.084***	0.034***
	(0.0040)	(0.0040)	(0.0074)			(0.0058)	(0.0039)
Age squared	-0.00069***	-0.00056***	-0.0011***			-0.0010***	-0.00038***
	(0.000049)	(0.000050)	(0.000100)			(0.000072)	(0.000047)
Constant	-0.0078**	0.0022	-0.032***	-0.0035	-0.000096	-0.0054	-0.0081**
	(0.0032)	(0.0030)	(0.0067)	(0.0040)	(0.0032)	(0.0045)	(0.0035)
	(0.024)	(0.027)	(0.058)	(0.024)	(0.033)	(0.028)	(0.033)
Industry FE	YES	YES	YES	YES	YES	YES	YES
$Quarter \times Year$ FE	YES	YES	YES	YES	YES	YES	YES
City FE	YES	YES	YES	YES	YES	YES	YES
Observations	123,437	96,614	26,823	71,372	52,065	62,670	60,767
R-squared	0.329	0.367	0.239	0.348	0.302	0.301	0.369

Notes: Standard errors clustered at the state level are shown in parentheses. Variables not shown include industry fixed effects, *quarter* × *year* fixed effects and city fixed effects. *** is significant at the 1% level; ** is significant at the 5% level; * is significant at the 10% level.

1.E Definitions of formal labor by country

Table 1.E.1: **Social protection (legalistic) definition of labor formality.**

Country	A worker is formal if she...
Argentina	has the right to a pension when retired
Bolivia (since 2000)	is affiliated with AFP (Administradora de Fondos de Pensiones)
Brazil	contributes to the Social Security system
Chile	is affiliated with any social security system
Colombia (ENH)	has the right to a pension when retired
Ecuador (ECV)	has the right to a pension when retired
El Salvador	is affiliated with any social security system (no information for domestic servants)
Guatemala	contributes to the IGSS (Instituto Guatemalteco de Seguridad Social)
Mexico (since 2000)	has the right to a pension when retired
Nicaragua	contributes to the INSS (Insituto Nicaraguense de Seguridad Social)
Paraguay	is affiliated with any social security system
Peru (since 1999)	is affiliated with any social security system
Uruguay (since 2001)	has the right to a pension when retired
Venezuela (2000-2003)	has the right to social benefits

Source: CEDLAC and Gasparini and Tornarolli (2007).

Bibliography

Albrecht, James, Lucas Navarro, and Susan Vroman (2009). "The Effects of Labour Market Policies in an Economy with an Informal Sector". In: *The Economic Journal* 119.539, pp. 1105–1129.

Alter Chen, Martha (2001). "Women and Informality: A Global Picture, the Global Movement". In: *SAIS Review* 21.1, pp. 71–82.

Anton, Arturo, Fausto Hernandez, and Santiago Levy (2013). *The end of informality in Mexico? Fiscal reform for universal social insurance*. New York: Inter-American Development Bank.

Arturo Anton Sarabia, Fausto Hernandez (2013). "VAT Collection and Social Security Contributions: Is There a Link?" In: *IDB Working Paper Series*. URL: http://www.iadb.org/en/publications/publication-detail,7101.html?id=68978.

Auten, Gerald and Robert Carroll (1999). "The Effect of Income Taxes on Household Income". In: *The Review of Economics and Statistics* 81.4, pp. 681–693.

Bernal Salazar, Raquel (2009). "The informal labor market in Colombia: identification and characterization". In: *Desarrollo y Sociedad* 63, pp. 145–208.

Bernardi, L. (2008). "Tax systems and tax reforms in Latin America". In: Routledge international studies in money and banking, 45. London; New York: Routledge. Chap. 11 MEXICO, pp. 414–450. ISBN: 0203481399 9780203481394. URL: http://www.crcnetbase.com/isbn/9780203481394.

Biles, James J. (2009). "Informal Work in Latin America: Competing Perspectives and Recent Debates". In: *Geography Compass* 3.1, pp. 214–236.

Boeri, Tito and Pietro Garibaldi (2006). "Shadow Sorting". In: *Fondazione Collegio Carlo Alberto Working Paper Series* 10.

Bosch, Mariano and Raymundo M. Campos-Vazquez (2014). "The Trade-Offs of Welfare Policies in Labor Markets with Informal Jobs: The Case of the Seguro Popular Program in Mexico". In: *American Economic Journal*.

Bosch, Mariano. et al. (2013). *Labor market effects of introducing unemployment benefits in an economy with high informality*. Washington, D.C.: Inter-American Development Bank, Research Dept.

Chant, Sylvia (1991). *Women and survival in Mexican cities: perspectives on gender, labour markets, and low-income households*. Manchester, UK: Manchester University Press.

Chen, Martha Alter (2007). "Rethinking the informal economy: linkages with the formal economy and the formal regulatory environment". In: *New York, NY United Nations, Dep. of Economic and Social Affairs* 46.

Correia, Elizabeth G. Katz; Maria Cecilia (2001). *The Economics of Gender in Mexico: Work, Family, State, and Market*. Washington DC: World Bank.

Dougherty, S. and O. Escobar (2013). "The Determinants of Informality in Mexico's States". In: *OECD Economics Department Working Papers* 1043.

Economic Co-operation, Organisation for and OECD Development (2011). *Taxing wages*. Paris: OECD Publishing.

Eissa, Nada and Jeffrey B. Liebman (1996). "Labor Supply Response to the Earned Income Tax Credit". In: *The Quarterly Journal of Economics* 111.2, pp. 605–637.

Fernandez, Andres and Felipe Meza (2014). "Informal Employment and Business Cycles in Emerging Economies: The Case of Mexico". In: *Review of Economic Dynamics*.

Fields, Gary S. (1975). "Rural-urban migration, urban employment and underemployment, and job-search activity in LDCs". In: *Journal of Development Economics* 2.2, pp. 165–187.

— (2009). "Segmented labor market models in developing countries". In: *The Oxford Handbook of Philosophy of Economics*.

Fiess, Norbert M., Marco Fugazza, and William F. Maloney (2010). "Informal self-employment and macroeconomic fluctuations". In: *Journal of Development Economics* 91.2, pp. 211–226.

Gasparini, Leonardo and Leopoldo Tornarolli (2007). "Labor Informality in Latin America and the Caribbean: Patterns and Trends from Household Survey Microdata". In: *Revista Desarrollo y Sociedad, Universidad de los Andes-CEDE.*

Gelber, A.M and J.W. Mitchell (2012). "Taxes and time allocation: Evidence from single women and men". In: *Review of Economic Studies* 79.3, pp. 863–897.

INEGI (2014). *Conociendo la base de datos de la ENOE: Datos ajustados a proyecciones de poblacion 2010*. Instituto Nacional de Estadistica y Geografia.

Jorge Alonso Ortiz, Julio C. Leal Ordo;ez (2011). "Taxes, Transfers and the Distribution of Employment in Mexico". In: *Munich Personal RePec Archive*.

Korlm, Ann-Sofie and Birthe Larsen (2004). "Does tax evasion affect unemployment and educational choice?" In: *Institute for Evaluation of Labour Market and Education Policy (IFAU) Working Paper Series* 4.

Levy, Santiago (2008). *Good Intentions, Bad Outcomes: Social Policy, Informality, and Economic Growth in Mexico*. Washington DC: Brookings Institution Press.

Maloney, William F. (1999). "Does Informality Imply Segmentation in Urban Labor Markets? Evidence from Sectoral Transitions in Mexico". In: *World Bank Economic Review* 13, pp. 275–302.

— (2004). "Informality Revisited". In: *World Development* 32.7, pp. 1159–1178.

Matias Busso Maria Victoria Fazio, Santiago Levy Algazi (2012). "(In)Formal and (Un) Productive: The Productivity Costs of Excessive Informality in Mexico". In: *IDB Working Paper Series* IDB-WP-341.

Mazumdar, Dipak (1976). "The Urban Informal Sector". In: *World Development* 4.8, pp. 655–679.

Meghir, Costas and David Phillips (2008). "Labour Supply and Taxes". In: *IZA discussion papers* 3405.

Meyer, Bruce D. and Dan T. Rosenbaum (2001). "Welfare, the Earned Income Tax Credit, and the Labor Supply of Single Mothers". In: *The Quarterly Journal of Economics* 116.3, pp. 1063–1114.

Ordoñez, Julio Cesar Leal (2014). "Tax collection, the informal sector, and productivity". In: *Review of Economic Dynamics* 17, pp. 262–286.

Pedro S. Amaral, Erwan Quintin (2006). "A competitive model of the informal sector". In: *Journal of Monetary Economics* 53.

Piketty, Thomas and Emmanuel Saez (2012). "Optimal Labor Income Taxation". In: *NBER Working Paper Series* 18521.

Schneider, Friedrich, Andreas Buehn, and Claudio E. Montenegro (2010). "Shadow Economies All over the World: New Estimates for 162 Countries from 1999 to 2007". In: *Policy Research Working Paper* 5356.

Schramm, Heidi R. (2014). "The Equilibrium Effects of Income Taxation on Formal and Informal Labor Markets". PhD thesis. University of Virginia.

Chapter 2

Latin American Aid Patterns: Democratic Political Regimes and US domestic Politics

Abstract

Is aid provision to Latin America different from the rest of the world? If so, why? In this paper we show how patterns of US foreign aid to Latin America not only differ from aid allocation observed elsewhere but it does so in manners that can potentially enhance current theories of aid provision. By using panel data on US aid, characteristics of recipient countries, and the economic views of US administrations from 1946 to 2001, we find that while political institutions and events in recipient countries greatly influence US aid allocations, the *ideological orientation* of US administrations can explain part of the divergent patterns of aid towards Latin America. Specifically, we find that while US liberal governments supported autocracies in Latin America, conservatives provided aid to countries with better democratic systems in the region. This finding calls for the inclusion of donor's domestic politics to account for unexplained variance in aid provision, which could influence aid allocation patterns to certain regions.

[0]Co-author: Jenny Guardado (email: jgr45@georgetown.edu)

2.1 Introduction

Are aid allocations to Latin America systematically "different" from the rest of the world? If so, why? In this paper we show how the pattern of US foreign aid to Latin America not only differ from patterns observed elsewhere but it does so in manners that could enrich current theories of aid provision. Based on the model developed by Bueno de Mesquita and Smith (2007) and Bueno de Mesquita and Smith (2009), we argue how this model, as applied to Latin America, can be enhanced by considering US domestic politics and not only its institutional framework. In particular, we build on Bueno de Mesquita and Smith (2007) and Bueno de Mesquita and Smith (2009) empirical analysis by adding a measure of ideological orientation. According to the authors, aid provision serves as a mechanism to obtain policies desirable to democracies in exchange for transfers to improve the survival of leaders from autocracies. As a result, a large determinant of aid is the democratic institutions of the recipient country, or the number of people whose support is necessary to keep the leader in office.

Yet, recent findings have shown how the ideological orientation of donor countries can explain foreign aid allocations (Milner and D. Tingley, 2010; Breuning, 1995; Fleck and Kilby, 2010; Goldstein and Moss, 2005). That is, whether governments are economically conservative or liberal influences the amount of aid given and the preferred recipients of such aid. Moreover, the ideological orientation of donor countries also influences the choice among different types of foreign policy available, such as economic aid or trade policies (Milner and D. Tingley, 2010). Therefore, ideology has been shown to play a substantial role in foreign policy.

In this paper we incorporate the literature on ideological orientation to the framework of Bueno de Mesquita and Smith (2007) and Bueno de Mesquita and Smith (2009). As explained in that model, the leaders of donor countries must see in its benefit the policy concessions obtained via aid. However, what is perceived as "beneficial" may vary from government to government within the same "institu-

tional" framework. For instance, leftist governments can gain supporters at home by providing large amounts of aid to developing countries in exchange for democratic progress, while conservatives might perceive aid to be ineffectual and rather support a broader trade policy to open markets for their domestic constituency's export products. In either case, ideological views might affect the desirability of aid versus other policy tools and explain the amount of aid given. Finally, among other goals, the paper provides further evidence[1] of how ideology explains US aid behavior towards Latin America.

This paper uses country panel data from various sources to analyze the patterns of aid allocation towards Latin America and the rest of the world. Three main findings are worth noticing: First, US aid provision to Latin America does follow a different pattern from that of the rest of the world. After controlling for numerous political, economic and geographical factors and contrary to popular explanations, it appears that the US did not always use aid as a transfer to autocracies in exchange for policy concessions in the region. However, with respect to the rest of the world, aid appears to be directed to more democratic countries, consistent with the results in Bueno de Mesquita and Smith (2007).

Second, we find that donor characteristics, such as the ideology of the US government, influenced the extent to which foreign aid was used to obtain policy concessions from Latin American countries in the 1946-2001 period. More interestingly, the ideology of the US government does not explain much of US aid provision to the rest of world but it does to Latin America. In addition, CIA interventions do not affect the amount of aid given to Latin American countries.

Our third finding shows that ideology matters for aid provision to Latin America. Conservative parties in the US increased the amount of aid provided to more democratic governments in the region. This finding contrasts with that for the rest of the world, where US ideology does not seem to have any effect on the amount

[1]By refining the analysis done in (Milner and D. Tingley, 2010) and Dustin Tingley (2010) to the context of Latin America and comparing it to the rest of the world.

of aid provision or on the decision to provide aid.

In this paper we show that features related to the political institutions of the recipient country (e.g. democracies) are important for Latin America. After controlling for changes in US domestic politics, particularly the orientation of the ruling party, such features can explain changes towards Latin American countries but less so to the rest of the world.

The paper adds to the current literature in three ways. First, most studies on US policy focus on US policy towards Latin America without contrasting with what is the norm in the rest of the world. Rather, US policy towards Latin America is assumed to be different as the sole justification of a regional approach. However, without a contrast of how aid is allocated elsewhere and controlling for multiple factors is hard to justify such approach. In this paper, we do not assume that Latin America is different *a priori*, but rather show with evidence the contrast in aid provision to the region with the patterns observed in the rest of the world.

Second, studies claim there exists an influence of US ideological orientation on foreign aid (Milner and D. Tingley, 2010). Yet, these have not been systematically tested within a broader model of aid provision. Moreover, these models have not looked at their regional applicability, especially to Latin America. The inclusion of ideological orientations would provide a more nuanced view to the more recent models of aid provision (e.g. Bueno de Mesquita and Smith, 2007 and Bueno de Mesquita and Smith, 2009). Results show that in some cases US ideology does matter, especially towards a region with high historical salience to the US such as Latin America. Future research will look into why the region as a whole warrants a different treatment, while in this paper we look solely at which factors mattered most from the US perspective and controlling for country (not regional) specific factors.

Finally, Berger et al. (2013) claim that CIA interventions led to an increase in foreign aid from the US. They look at the value of US economic and military aid

received by each country, and test whether CIA interventions led to an increase in US foreign aid. Following their specification we find that for Latin America, CIA interventions are not relevant for the amount of aid or the probability of receiving aid, while for the rest of the world, it seems that it is a determinant of aid provision from the US. Our results confirm that Latin America has been a region that the US has treated differently from the rest of the world in terms of aid provision, where ideology of US political parties and political regimes of receiving countries are more important in the determination of how much aid is provided to the region.

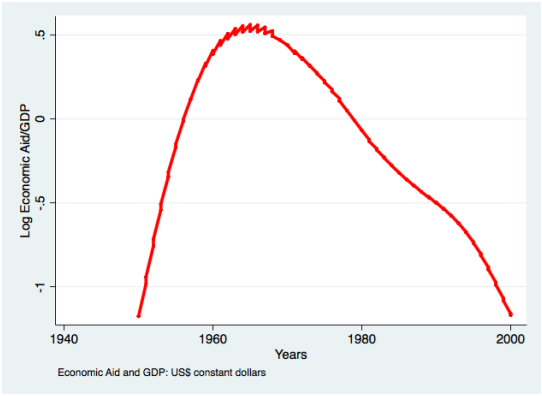
The study has a regional outlook to unravel variables that might be overlooked in a cross-country analysis. For instance, it is well known that the strategic interest of the US in Egypt and Israel to ensure stability in the Middle East has been followed by a large amount of aid to both countries. However, Figure 2.1 shows a particular pattern of aid allocation to Latin America in which the levels of economic aid have actually followed an upward slope while in the rest of the world aid has been declining.[2] For instance, prior to the oil crisis (the height of US aid) Latin America was receiving an amount of aid similar in quantity to that of the rest of the world combined. Yet, after 1973 Latin America seems to be following a downward slope similar to the rest of the world. In contrast, Figure 2.2 exhibits the pattern of military aid towards Latin America, which has seen a similar trend just as elsewhere.

2.2 Literature Review

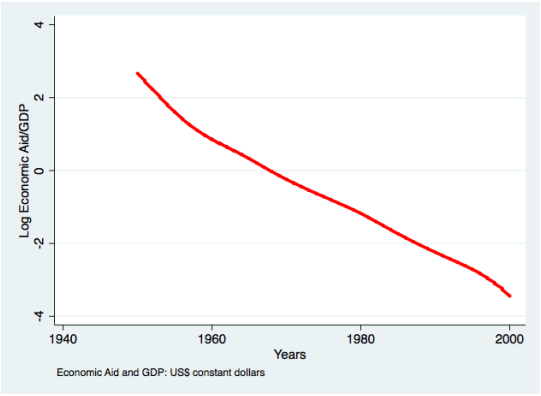
The role of foreign aid as an instrument of foreign policy is a well-established fact. The type of aid, amount, and recipients are said to be influenced by an array of geopolitical and strategic factors in both donor and recipient countries. Explanations based on features of the international arena argue that foreign aid allocations

[2]Sub-Saharan Africa is the only other region that shows a different pattern compared to the rest of the world. Our analysis will only focus in Latin America

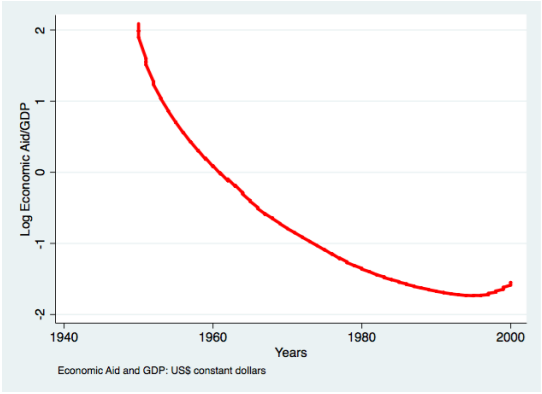
Figure 2.1: **Economic Aid (1946-2001)**



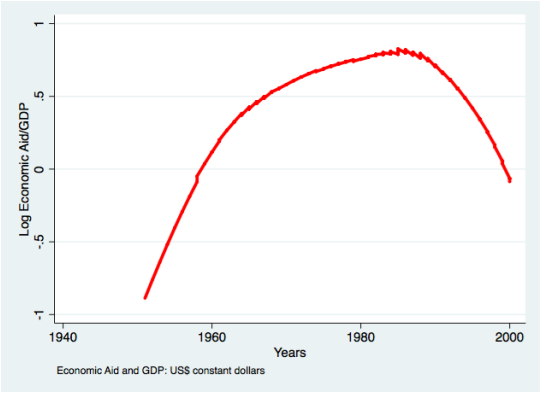
(a) Latin America Economic Aid



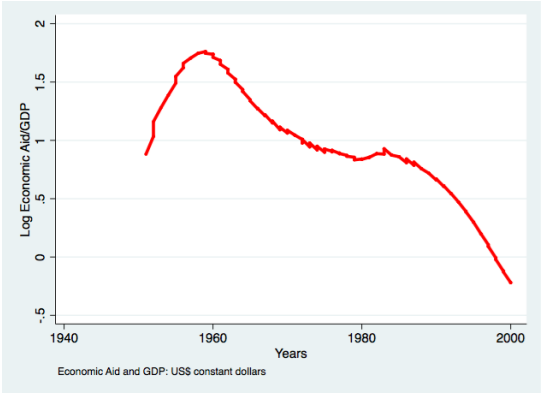
(b) East Asia and Pacific Economic Aid



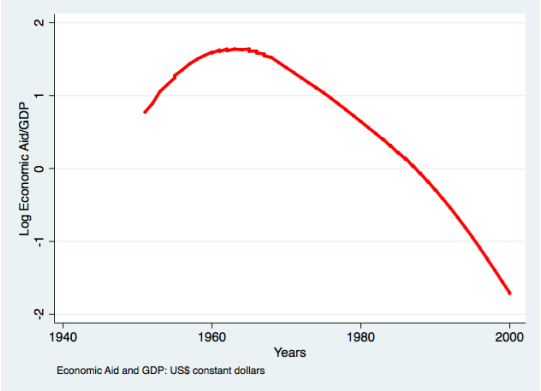
(c) Europe and Central Asia Economic Aid



(d) Sub-Saharan Africa



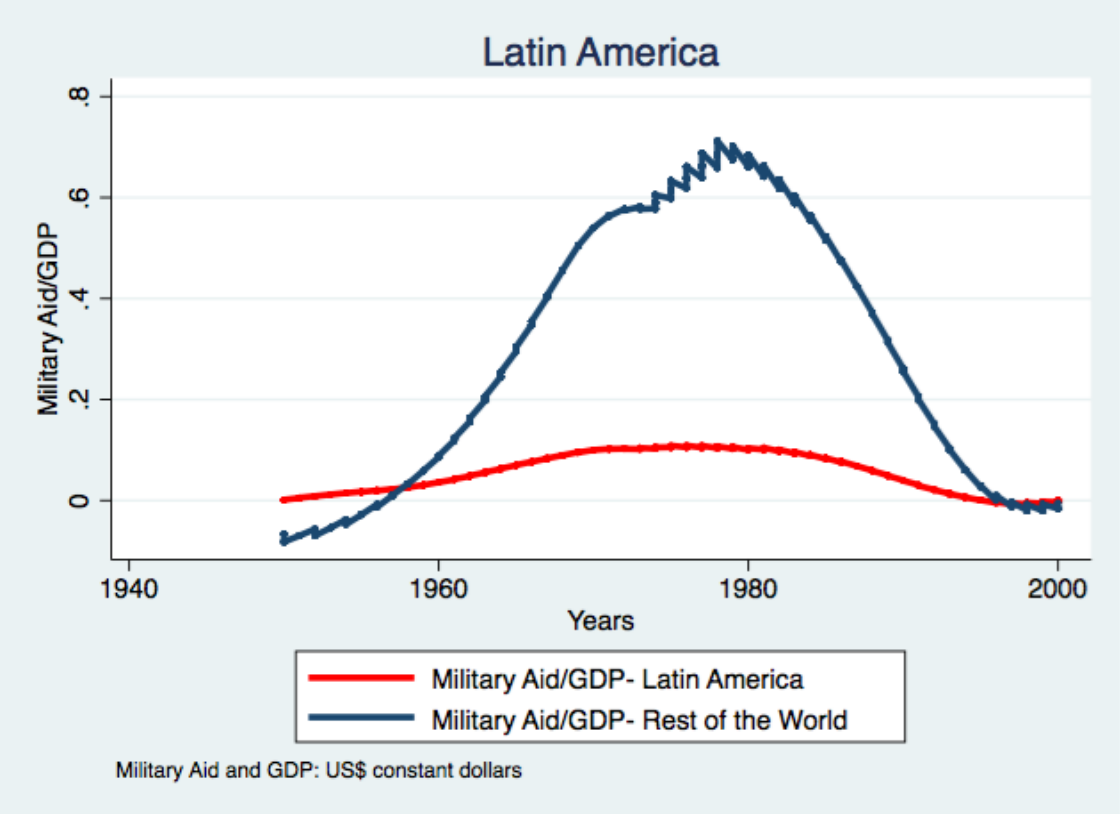
(e) Middel East and North Africa Economic Aid)



(f) South Asia Economic Aid

would be primarily driven by security considerations, especially for hegemonic countries like the US. In such cases, foreign aid (especially military aid) might reflect security driven relations such as alliances (Bueno de Mesquita, 1981). In contrast, other explanations emphasize the importance of economic goals, specifically trade,

Figure 2.2: Military Aid Latin America and the World: 1946-2001



to account for the variation in aid provision. Export groups from donor countries might benefit from additional markets to offer their products, while import groups and governments of recipient countries might also benefit from an increased amount of trade by obtaining higher revenue from tariffs. Therefore, donor governments may be pressured to use aid to obtain trade policy concessions from certain countries (Bueno de Mesquita and Smith, 2007) or simply to reward major trading partners (Meernik, Krueger, and Poe, 1998).

Others argue that foreign aid aims to achieve convergence in ideological terms between donors and recipients. In the case of the US, the implication is that aid is given for the promotion of democracy and human rights, which during the cold war was perceived as an anti-communist goal (Packenham, 1973). The promotion of democracy and human rights also had a security aspect, since it would help defuse conflict with other democracies (Maoz and Russett, 1993). Moreover, empirical

evidence supports the "promotion" of democracy effect. For instance, Alesina and Dollar (2000) find that countries going through a democratization episode receive an immediate increase in aid. Also, Bueno de Mesquita and Smith (2007) observe that once aid is given to autocratic institutions, further increases in democratization lead to increases in foreign aid allocation. Yet, the authors also find that other political and strategic considerations play a role in explaining US aid allocation; international donors are influenced by policy considerations, past colonial history and prevailing economic conditions of the recipient countries.

A drawback with empirical studies on aid provision is the lack of a general explanation of why the US would target aid to improve democracy and trade openness and why the recipient countries would accept. The exception is Bueno de Mesquita and Smith (2007), who focus on the incentives of both donor and recipient countries and argue that aid allocation is given to regimes colloquially known as "rigged electoral autocracies". The reason is that autocracies would reduce the costs of obtaining a policy concession from the recipient state, while the politics of the donor country benefit in terms of policy concessions from the recipient.

The difficulty with the Bueno de Mesquita and Smith (2007) and Bueno de Mesquita and Smith (2009) model is that "institutions", while explaining a lot of the observed variation, does not account for the "preferences" or ideological orientation of the group in power. Such preferences are important, considering they influence both the decision to choose aid as a policy (instead of trade, for example) and the required "concessions" from recipient countries. Yet, the empirical evidence of the effect of ideology on foreign aid allocations to the World is mixed at best (Breuning, 1995). Some studies have found a role for ideology when looking at certain regions. For instance, (Goldstein and Moss, 2005) find that Republican governments provided more aid to Africa than Democrats did. Thus suggesting an important role of ideology on regional allocations. Some authors argue that the reasons of such differences lie in that conservative governments are driven by

more commercial interests (e.g. broadening trade) while liberal governments tend to allocate aid based on development concerns (Fleck and Kilby, 2010). Therefore, we should observe a larger emphasis on the resources and the amount of trade under different US administrations. In this paper we incorporate ideological considerations to enhance the explanatory power of the empirical model.

Among the studies looking at US - Latin America relations, foreign aid is frequently addressed within the framework of human rights promotion. Among the noteworthy studies of the region are: First, Cingranelli and Pasquarello (1985) who revisit the claim that US foreign aid tended to reward human right violators and punish those countries with a good record. In their view, human rights considerations have been increasingly important in bilateral aid allocations to the region since the mid 1970s. Yet, this finding has been challenged by McCormick and Mitchell (1988) who argue that the relationship is non-existent once accounting for some of the countries' measures of human rights violations. Finally, Meernik, Krueger, and Poe (1998) revisit the whole impact of human rights and find that although human rights concerns were not the major factor, they did play a role in accounting for aid allocation. They argue that the increasing emphasis on human rights protection has been driven by the influence of US Congress rather than by security concerns of the US President. A well known example is that of aid support directed to the Nicaraguan Contras and Angolan UNITA promoted by the US President which Congress later amended (Lagon, 1992). Thus, the role of US aid in promoting (or hindering) human rights remains an open question.

A related issue is the role played by the US in the rise of authoritarianism in Latin America via the support of military juntas during the 1960s and 1970s. Muller (1985) argues that the breakdown of democracy in the Third World (hence Latin America) is a by-product of the competition between the US and Soviet Union, in which military aid, rise of leftist forces and anti-communist policy hindered democracy in certain regions. Based on this claim, Sanchez (2003) finds that the

US provided military assistance to military governments friendly to US interests, thus contributing to the breakdown of democracy in the region. Yet, the latter study only focuses on military aid, therefore excluding the role of economic aid as well as more rigorous tests assessing the impact of military juntas by controlling for geographical, economic and political factors. Therefore, we analyze the impact of military aid on the government regime to see whether the correlation found holds once controlling for alternative accounts.

Finally, a different strand of the literature looks at whether the US provided a larger amount of military assistance to countries facing a strong leftist insurgent threat as was the case in many Latin American countries. This policy is the so called "Reagan Doctrine" which explicitly aided anti-communist guerrillas in the Third World as in Afghanistan, Nicaragua, Angola and Cambodia (Lagon, 1992). Yet, such approach might not only be limited to Reagan, since the US provided aid to military groups to overthrow the Arbenz government in Guatemala (1954) and invade Cuba in 1961. Therefore, it is possible that this policy might have started well before the Reagan administration and it is based on the cold war constraints of not being able to directly challenge the Soviet Union (Lagon, 1992). Throughout the analysis we include a control for the possibility that the ideological orientation of certain presidents might matter more than others (e.g. the Reagan period).

Beyond geographical variation in US patterns of aid allocation, the literature has also recognized variation across time in US aid policies. In this paper we will argue not only that there is geographic variation in terms of aid allocation policy in the world, as seen in Figure 2.1 and Figure 2.2. We attempt to explain divergent patterns towards Latin America, but also clarify misconceptions in US foreign policy (as pertaining aid) to the region. In sum, to the best of our knowledge there is no systematic study of US-Latin America relations which has directly addressed the issue of US ideological orientation on aid provision.

2.3 Data

The data utilized in this paper was gathered from four main sources: Bueno de Mesquita and Smith (2007), Milner and D. Tingley (2010), Berger et al. (2013) and the COW trade dataset.

Aid: The main dependent variable in this study is the logarithm of economic and military aid ($lnEaid$ and $lnMaid$) provided to Latin America and the World in constant (1996) U.S. dollars as reported by the U.S. Agency for International Development (USAID) from 1946 to 2001. However, unlike Bueno de Mesquita and Smith (2007) we conduct a separate analysis to economic and military aid but use total economic aid ($lnTaid$) as a robustness check. In addition to the amount of aid given, the analysis also accounts for the possibility of receiving any amount of aid: ($Eaid$, $Maid$, $Taid$) Economic, Military and Total aid, respectively. Another robustness check comes from the OECD/DAC database on aid commitments. Specifically, we look at total aid commitments including both bilateral and multilateral aid.[3]

Political Variables: Coalition size (W) is an index between 0 and 1 with higher numbers measuring the degree to which a leader is bound by the ruling coalition. The selectorate size (eS) index goes from 0 to 1 and measures the proportion of people with a potential influence over the decision of who is going to be the leader. Following Bueno de Mesquita and Smith (2007) the selectorate size is reparamatrized as $S * (1 - W)$ to emphasize the importance of selectorate size within autocracies. Both variables are measured for the period 1946-2001.

The main independent political variable is the US government ideological orientation ($Ideology$) obtained from Milner and D. Tingley (2010)[4] based on the Comparative Manifestos Project. Their idea was to code the manifestos of politi-

[3] A drawback from the latter is that the period in which this variable is available goes from 1960-2001.

[4] Tingley's (2010) is based on a large number of donors (all developed democracies) from which we only focus on the ideological orientation of the US.

cal parties, in particular the economic orientation of the government in place using factor analysis to obtain the first dimension of the economic features which then allows to calculate the place of each party with regression scores.[5] These procedures lead to a continuous measure of ideology based on the economic positions of parties in power. The idea is that aid and trade policy might be as influenced by the positions of parties within governments as they are by party positions outside government (Milner and D. Tingley, 2010). The inclusion of the ideological variable provides some variation in the degree of consensus and policy preferences over time prevailing in US domestic politics which contrasts with the time-invariant measures of coalition size (or institutions). If we were to observe that changes in aid provision are not related to the ideological orientation of the US government then domestic politics by themselves cannot explain variations in aid provision to Latin America and the World. Finally, a limitation with this variable is that it is only coded starting 1960, therefore the analysis using these variables explains a shorter time-period, that of 1960-2001.

We also include as political variable a measure of successful CIA interventions (from Berger et al., 2013). This indicator equals one if the CIA either installed a foreign leader or provided covert support for the regime once in power. The idea is that aid might be influenced by CIA interventions, which in their paper would lead to an increase in US imports. This variable provides evidence of the increased political influence arising from CIA interventions in receiving countries, which could provide a potential explanation of the amount of aid provided or the probability of aid provision.

Another political variable relates to the recipient's domestic politics: the extent of political unrest based on Bank's (2012) measure of political conflict as a weighted measure of different conflict instances such as assassinations, strikes, guerrilla, purges, riots, revolutions and protests as well as an indicator for civil war

[5]For more details, consult Milner and D. Tingley (2010).

presence. Yet, given the political situation of Latin America, we will separately use the coding of guerrilla instances to add the relative strength of such insurgency movements. Finally, we use Bueno de Mesquita (1981) measure of alignment with the United States based on security alliance portfolios (*tau*). As suggested by the literature, security concerns might be driving the provision of aid during the cold-war period: more aid given to "allies" than otherwise.

Table 2.1: **Summary statistics for Latin America and the World**

	Rest of the World		Latin America	
Variable	Mean	Obs.	Mean	Obs.
lnEconomicaid (ln US $ constant)	9.972	3,443	9.558	1,228
	(2.111)		(1.909)	
Economicaid (US $ constant)	0.472	7,294	0.663	1,853
	(0.499)		(0.473)	
US intervention (Dummy 0-1-CIA successful interv.)	0.159	4,259	0.363	1,071
	(0.365)		(0.481)	
Political Regime (Index (autocratic)0-1(democratic))	0.549	6,725	0.579	1,755
	(0.293)		(0.246)	
Ideology (Index (liberal)0-10(conservative))	5.677	5,453	5.677	1,384
	(0.366)		(0.366)	
$W * Ideology$	3.104	5,159	3.379	1,320
	(1.732)		(1.515)	
$W * Intervention$	0.063	4,240	0.148	1,071
	(0.182)		(0.244)	
Selectorate size (Index 0-1: proportion people potential influence)	0.299	5,417	0.296	1,400
	(0.272)		(0.227)	
Share of government GDP (US$ constant)	12.897	4,163	12.42	1,306
	(1.894)		(1.748)	
Ln per capita Income (US$ constant)	7.644	4,620	7.999	999
	(1.096)		(0.574)	

Standard deviation in parentheses.

Economic variables: The two most important economic variables to account for in this paper are the resources of the recipient and donor countries as well as the amount and direction of trade between both countries. For the latter we use COW trade data which was obtained from the IMF's Direction of Trade Statistics. We specifically use the flow of exports (USexports) from the US to the world as well as the flow of imports (USimports) into the US. This measure is an improvement over a combined measure of exports and imports, since foreign aid can be used to purchase imports creating simultaneity problems. Following Meernik, Krueger, and Poe (1998) we use the value of US imports from each country (or the amount

of exports to the US).

As a robustness check we also use the logarithm of bilateral trade between two countries in 1996 US constant dollars. To measure the level of resources of recipient countries we follow Bueno de Mesquita and Smith (2007) and use the share of government GDP multiplied by the lagged logarithm of GDP of the recipient country (RB_{t-1}). Similarly, the level of resources for the donor country is measured as the share of GDP (RA) multiplied by US GDP. Following Bueno de Mesquita and Smith (2007) we also include the squared version of RB to reflect the non-monotonicity anticipated in the effect of RB. Since the term RB already includes per capita income and population, which belong to the econometric specification in their own right (as a control for poverty and humanitarian need), we exclude them from our specification and only keep RB and RB^2. All economic variables are available from 1946 to 2001.

For the US, the salience of the policy concession sought from Bueno de Mesquita and Smith (2007) is measured with three indicators: *distance*, *population*, and *colony*. *distance* is estimated as the logarithm of the distance in miles between each prospective recipient's and each prospective OECD donor's capital city. *population* is measured as the logarithm of the prospective recipient countries population in millions as reported by Penn World Tables. *colony* is a dummy variable coded as 1 if the potential recipient country had been a colony of the prospective donor. The general idea is that policy concessions from geographically closer, more populous countries are valued more than comparable concessions from small distant countries. Similarly, former colonies hold higher salience for donors than do states with which they had no special prior relationship.

The prediction is that the role of institutions, resources, CIA interventions, and policy salience should not be different in Latin America. However, as already seen in the introduction, different patterns of US aid to Latin America are noticeable. If such variation can be explained with institutional variables in the way predicted

by Bueno de Mesquita and Smith (2007) there is stronger evidence for their model. If not, other hypotheses might have some explanatory power in Latin America than in the rest of the world in which case an alternative hypothesis will be tested following the literature. Although this runs against Bueno de Mesquita and Smith (2007) and Bueno de Mesquita and Smith (2009) it does mean there are special circumstances which make the predictions of their model less evident. Among these "special circumstances" we particularly account for the role ideology plays in explaining the results observed in Latin America, as noticed by Dunning (2004) as well as the role of CIA interventions as shown in Berger et al. (2013). Also, it is possible that a confounder factor, such as guerrilla warfare, might be explaining the results observed in the region, thus we carefully look at this possibility.

2.4 Methodology

To explore these hypotheses, we look at aid provision patterns from 1946 to 2001. The analysis can then elucidate to what extent are democratic institutions[6] different across time and different geographic regions. Below we carry out two sets of estimations: one uses a fixed effects model for all US aid recipients between 1946 and 2001 and the other looks at the probability of aid provision. The fixed effects model can be written as:

$$lnAid_{it} = \alpha_i + \beta W + \delta USInterv_t + \lambda W * Ideo + \mathbf{X}_{it-1} + \mathbf{M}_{it} + \gamma_t + \varepsilon_{it} \quad (2.1)$$

where i is the country, t is the year, W is the type of political regime (1=democracy, 0=autocracy), $USIntervention_t$ is an indicator variable that equals one if the CIA either installed a foreign leader or provided covert support for the regime in country i at time t; α_i are country fixed effects; γ_t are time fixed effects; $W * Ideo$ is the type of political regime multiplied by the score of US economic ideology prevailing

[6] As in the Bueno de Mesquita and Smith (2007) model.

in year t (conservative or liberal); $lnAid$ is the amount of aid given, $\mathbf{X}$ is a vector of explanatory variables lagged by one year (e.g. recipient resources), and $\mathbf{M}$ are explanatory variables not lagged.

This model has the purpose of looking at whether the levels of the explanatory variables are correlated with the level of the dependent variable controlling for inherent country heterogeneity. Also, the inclusion of fixed effects allows for trends in aid to vary from country to country and controls for omitted variable bias. Yet, looking at within country variation precludes the analysis the comparison of countries who do not receive any aid at all. Therefore, we also estimate the probability of any amount of aid being given $Pr(aid = 1)$:

$$Pr(Aid_{it} = 1) = \alpha + \beta W + \delta USInterv_t + \lambda W * Ideo + \mathbf{X}_{it-1} + \mathbf{M}_{it} + \gamma_t + \varepsilon_{it} \quad (2.2)$$

where all explanatory variables are defined as mentioned above. The coefficient of interest in both cases is β, which according to Bueno de Mesquita and Smith (2007) should be negative in Equation 2.1 and positive in Equation 2.2.

In addition to the models estimated, the samples observe patterns overlooked in cross-country analysis. The samples used will be specified in every table.

2.5 Results

The mechanism proposed by Bueno de Mesquita and Smith (2007) highlights the role of institutions, the importance of policy and the resources of the recipient state as determinants of aid. Such explanation clearly differs from the altruistic motivations argued by others (Lumsdaine, 1993). Moreover, the authors provide a coherent explanation as to why we should observe aid giving behavior and in what amount, which is an improvement on previous literature. However, it appears that this mechanism better explains some regions than others. Table 2.2 (below) shows Model 1 and Model 4 of Bueno de Mesquita and Smith (2007) Tables 1 and 4,

including also CIA interventions from the US.

The dependent variable of the first two columns is the amount of aid given to the world and to Latin America in particular (when we multiply the effect by the region), while columns (3) and (4) have as a dependent variable a binary outcome (aid or no aid) and it uses a logit model.[7] To classify countries into regions we followed the World Bank classification (see Appendix 2.B for classification). For reasons of space we only present and discuss the main independent variables discussed in the previous sections (political and economic).[8]

Table 2.2: **Aid Provision to Latin America and the World: 1946-2001**

	$lnEconomicAid$		$Pr(aid)$	
	(1)	(2)	(3)	(4)
US intervention (CIA intervention dummy)		0.899***		0.212**
		(0.272)		(0.087)
$USintervention * LAC$ (Dummy)		-0.606		-0.142
		(0.417)		(0.097)
Political Regime (W) (Index 0-1 democratic)	0.506	0.640*	-0.023	0.021
	(0.398)	(0.374)	(0.070)	(0.075)
$PoliticalRegime * LAC$	0.165	0.166	-0.036	-0.043
	(0.551)	(0.532)	(0.077)	(0.086)
Selectorate size (eS) (Index 0-1 influence)	-0.052	-0.098	-0.020	-0.034
	(0.228)	(0.218)	(0.036)	(0.036)
Share of government GDP (RB)	4.046***	4.061***	0.048	0.027
	(0.895)	(0.870)	(0.207)	(0.198)
RB^2	-0.188***	-0.187***	-0.006	-0.005
	(0.036)	(0.035)	(0.008)	(0.007)
Country fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Observations	2,525	2,525	3,536	3,536
R-squared	0.169	0.178	0.098	0.113

Controls not shown: Ln pop, alignment with US (tau), Civil war, Soviet Influence, amount of trade.
Notes: Robust standard errors clustered at the country level are shown in parentheses. *** is significant at the 1% level; ** is significant at the 5% level; * is significant at the 10% level.

Columns (1) and (2) indicate that similarly to Bueno de Mesquita and Smith's (2007) results, the overall patterns of the amount of aid given to the world are not significant. The probability of receiving economic aid is not granted differently

[7] The results are also robust to other binary outcomes models (e.g. probit).
[8] Complete results are available upon request.

to governments with democracies than to less democratic governments, even after excluding Latin America. However, as noticed in Columns (2) and (4), with the inclusion of CIA interventions, the type of political regime has an effect for the world, although not for Latin America. That is, more democratic countries receive more aid.

For military aid (see Appendix Table 2.A.1) results are similar to Bueno de Mesquita and Smith (2007), where more aid is given to more autocratic governments in Latin America, while for the rest of the world, military aid is given to more democratic countries. The probability of receiving aid is only significant for Latin America, with a positive coefficient, while for the rest of the world is not significant.

An alternative is that US aid policy not only varied geographically but also by CIA interventions (as Berger et al., 2013 provide for the world). Similar to Berger et al. (2013) findings, we also show that the amount of aid increased following CIA interventions for the world, however this is not the case for Latin America, where at most we can say that the coefficient is not different from zero.

So far, we have just established regional differences in US aid allocation, yet the implications are potentially important: First, these results show that there are certain regions for which the exchange of aid-for-policy concessions for dictatorships are more applicable than others. Thus, the US may not always target dictatorships to pursue its preferred policies, at least in Latin America. Second, this result seems peculiar to Latin America since for other regions in the sample (e.g. Sub-Saharan Africa), when able to be consistently estimated, there is no difference.[9]

In the next section we analyze in depth two factors which would, in theory, enhance the reliance of the US on autocratic governments to pursue its policy aims. First, the possibility that throughout the period, the ideological orientation of US

[9]Results for Sub-Saharan Africa suggest no effect of the political regime. In the Middle East, South Asia and East Asia, the lack of enough observations prevents us from drawing any definite conclusion. Although alternative models could be estimated for which enough observations exist, the lack of all the controls would make the results less comparable across specifications.

administrations and CIA interventions would explain the amount of aid and trade directed towards receiving countries. Second, we will then account for domestic events (such as Guerrillas and Civil War) that could explain the divergent patterns observed in Latin America and the rest of the world.

2.5.1 Can US domestic ideology explain Aid policies?

According to the approach by Bueno de Mesquita and Smith (2007), democratic donors have a higher incentive to engage in policy-for-aid deals. As noticed by the authors, the problem of testing this claim empirically is that there is little variation in the donors democratic governments (W) throughout the period analyzed since it is already a consolidated democracy. An alternative explanation is that the temporal shift in foreign aid observed in Latin America is driven by changes in US domestic politics. For instance, Milner and D. Tingley (2010) show that aid and trade policy can not always be used as a substitute of the other, even if the foreign policy goal is the same. The reason is that these policies tend to be supported (and disliked) by different constituency groups in the US. Therefore, it is not uncommon to see different US policy instruments to move in different directions even if the policy goal is similar. Thus, if there is a change in the domestic interests of the US towards protecting trade then we should see a surge in foreign aid and vice-versa. Moreover, it has been shown in cross-country regression analysis of donors' aid provision that the ideology of the government in turn can affect aid preferences with right wing governments severely cutting the provision of aid (Milner and D. Tingley, 2010). Therefore, instead of looking at the political regime of the donor country, we look at varying levels of support and ideology within the given institutions for aid levels. The data is collected from Milner and D. Tingley (2010), from which we use as variables the levels of ideology and their level of support in US Congress (1960-2001) weighted by voting levels and all other variables that were included previously.

Table 2.3: **US Ideology on Aid Provision to Latin America and the World: 1960-2001**

	$lnEconomicAid$ (1)	$Pr(aid)$ (2)
US intervention	0.969**	0.106**
	(0.407)	(0.045)
$USintervention * LAC$	-1.245**	-0.118*
	(0.605)	(0.065)
Political Regime (W)(Index 0-1 democratic)	0.572	-0.164
	(2.309)	(0.218)
$PoliticalRegime * LAC$	-9.080	-0.236
	(5.622)	(0.490)
Ideology (Index 0-10 conservative)	0.233	0.003
	(0.844)	(0.027)
$Ideology * LAC$	-1.408**	-0.060
	(0.620)	(0.056)
$W * Ideo$	0.017	0.025
	(0.374)	(0.038)
$W * Ideo * LAC$	1.573	0.040
	(0.994)	(0.085)
$W * Intervention$	-0.452	0.197***
	(1.047)	(0.076)
$W * Intervention * LAC$	1.386	-0.092
	(1.180)	(0.113)
Selectorate size (eS) (Index 0-1 influence)	0.081	0.000
	(0.208)	(0.023)
Share of government GDP (RB)	4.565***	0.117
	(1.164)	(0.073)
RB^2	-0.210***	-0.010***
	(0.048)	(0.003)
Country fixed effects	Yes	Yes
Year fixed effects	Yes	Yes
Observations	2,137	3,072
R-squared	0.199	0.056

Controls not shown: Ln pop, alignment with US (tau), Civil war, Soviet Influence, amount of trade, new leader and years in office.

Notes: Robust standard errors clustered at the country level are shown in parentheses. *** is significant at the 1% level; ** is significant at the 5% level; * is significant at the 10% level.

Column (1) in Table 2.3 estimates the effect of US partisanship on US economic aid to the rest of the world, and specifically to Latin America. As noticed, there is no effect of ideology or political regime on the amount of aid given conditional on some aid being given to the rest of the world. We find a negative and significant effect of the ideology on the amount of aid given in Column (1) which includes the interaction term of Ideology and Latin America (LAC). This result is similar to Milner and D. Tingley (2010) who finds that a more conservative stance reduces aid provision in a cross-country analysis of aid-donors, yet, the same data also reveals that looking into the observed variation in US aid giving cannot be explained in purely ideological terms.[10] Therefore, the results obtained in Column (1) are quite interesting since we observe how for Latin America, US aid provision has been partly influenced by the ideological movements in the US which differ from the rest of world as seen both in Milner and D. Tingley (2010) data and our analysis of Bueno de Mesquita and Smith (2007) and Berger et al. (2013)[11]

In terms of explaining CIA interventions, we also find differences in the pattern comparing the rest of the world to Latin America. While our results for the rest of the world are similar to the ones obtained by Berger et al. (2013), it appears that CIA interventions in Latin America have the opposite effect. Column (1) suggests that aid provision to Latin America was influenced by the CIA interventions as well as the ideological nature of the US government.

By looking at the probability of receiving economic aid, we conclude that the likelihood of receiving aid does not depend on either the ideology or political regime in the region. However, CIA interventions in the region do seem to have an effect in Latin America (see Table 2.3).

However, when we focus on military aid, the amount of aid to the rest of the

[10]This analysis was conducted with the author's (Milner and D. Tingley, 2010) own replication data. Their analysis does not separate between aid given by country, therefore it cannot be disaggregated to account for Latin America. Once we do so, the results prove to be different from the overall pattern exhibited.

[11]To account for the possibility of autocorrelation we include newey-west robust standard errors in Table 2.A.4 in the Appendix.

world decreases with more conservative US governments, but has no significant effect for Latin America. While probability of receiving aid in Latin America decreases with more conservative US governments (see Appendix Table 2.A.2).

Looking at the descriptive statistics in Table 2.1 in section 2.3 we can see that ideology levels were around 5.67, with higher values indicating more conservative stances. This suggests that the presence of a higher level (on average) of conservative ideology would be associated with a larger provision of aid to Latin America.

Is economic aid also affected by exports? In Table 2.4 we include the interaction of the ideological variable with the political regime to see how it affects the amount of aid controlling for US exports to receiving countries. As seen in Column (1) the positive effect of the interaction ($W * Ideology$) remains for LAC, while for the rest of the world remains not significative. Second, we see that aid and trade follow a inverse relation for Latin America, while for the rest of the world, if the amount of aid increases, so does the amount of exports. However, US exports are not associated with the likelihood of giving economic aid for Latin America (Column (2)).

The results in Table 2.4 complement those presented in Table 2.3 by including trade as an explanatory variable. We find that a larger amount of aid comes along with higher access to the US market, in the rest of the world. This finding is consistent with Meernik, Krueger, and Poe (1998), who find a positive relationship between aid and trade, but does not hold for Latin America.

It is also worth noticing how ideology does not have an effect on the amount of aid for the rest of the world, but is highly significant for Latin America, even after controlling for US Exports, which points to the robustness of this particular finding.

Table 2.4: **US Ideology and Trade on Aid Provision to Latin America and the World: 1960-2001**

	$lnEconomicAid$ (1)	$Pr(aid)$ (2)
US intervention (Dummy)	1.185***	0.048
	(0.244)	(0.046)
$USintervention * LAC$	-1.368***	-0.066
	(0.326)	(0.064)
Political Regime (W) (Index 0-1democratic)	0.093	-0.203
	(1.468)	(0.215)
$PoliticalRegime * LAC$	-10.090***	-0.146
	(2.616)	(0.482)
Ideology (Index 0-10 conservative)	-0.350	-0.023
	(0.378)	(0.027)
$Ideology * LAC$	-1.052***	-0.041
	(0.284)	(0.055)
$W * Ideo$	0.104	0.030
	(0.256)	(0.037)
$W * Ideo * LAC$	1.750***	0.027
	(0.457)	(0.084)
$W * Intervention$	-0.257	0.335***
	(0.459)	(0.075)
$W * Intervention * LAC$	1.041*	-0.216*
	(0.613)	(0.112)
US Exports	0.423***	0.054***
	(0.060)	(0.009)
$USExports * LAC$	-0.431***	0.001
	(0.077)	(0.015)
Selectorate size (eS) (Index 0-1 influence)	0.083	-0.011
	(0.120)	(0.024)
Share of government GDP (RB)	4.315***	0.083
	(0.411)	(0.077)
RB^2	-0.203***	-0.008***
	(0.016)	(0.003)
Country fixed effects	Yes	Yes
Year fixed effects	Yes	Yes
Observations	2,098	2,930
R-squared	0.224	0.071

Controls not shown: Ln pop, alignment with US (tau), Civil war, Soviet Influence, amount of trade, new leader and years in office.

Notes: Robust standard errors clustered at the country level are shown in parentheses. *** is significant at the 1% level; ** is significant at the 5% level; * is significant at the 10% level.

2.5.2 Alternative Accounts: Guerrillas and Civil Wars

A different explanation to the divergent patterns observed might not be due to changes in US domestic politics or to changes in the size of its ruling coalitions. Rather, different aid patterns could be due to domestic events among recipient countries. That is, the emergence of domestic threats to US interests in autocratic governments may have induced a switch in US policy which was not present before. For instance, the so called "Reagan Doctrine" was a response to the emergence of Soviet-friendly guerrilla movements in the Third World (Lagon, 1992). Therefore, if Latin America saw a disproportionate emergence of strong pro-communist guerrilla movements in autocratic countries, then aid provision may have appeared targeted to autocratic governments when actually it was targeted towards counterinsurgency efforts.

In this section we test whether aid provision is affected by the presence of guerrilla forces or civil war. Previously we have followed the specifications used in Bueno de Mesquita and Smith (2007) and Berger et al. (2013) for the sake of deriving relevant comparisons thus including the indicator of civil war presence in the country. Although the presence of civil war is relevant in the world context, in particular sub-Saharan Africa, its low presence in Latin America might have not captured the specific type of threat posed to US interests in this region.

Although many of these conflicts are originated by the presence of guerrilla groups, the lack of full-scale civil wars in the region calls for a more nuanced definition of conflict for the case. Moreover, in the tables presented previously, civil war has a small or null effect on the probability of receiving aid, and only in the specifications in which Latin America is excluded (as expected). A different case is its influence on exports to the US in which the presence of civil war was associated with a larger number of exports. The explanation lies in the nature of US interventions in civil war, since it had a larger emphasis on securing supplies to the US. In Table 2.5, we do not find an effect of guerrillas on the probability

of receiving aid (column 2) or on the amount of aid received (column 1) in Latin America and the world. However, Reagan has a negative and significant effect for LAC, while for the rest of the world, the effect is not significant.

Table 2.A.3 in the Appendix provides evidence on how the presence of guerrilla warfare is negatively related to the amount of military aid in the world but has not significant effect in Latin America. Furthermore, the probability of aid is not related to the presence of guerrilla in the region. This assertion runs against Sanchez (2003) arguing that military aid has contributed to democratic breakdown in the region, via its targeted aid to military juntas in the presence of insurgency threats.

As mentioned before, the presence of guerrilla threats does not have an effect on the amount of military aid given, conditional on any aid being given. This finding runs against the apparent influence of the Reagan Doctrine and the fact that it might be targeted for counterinsurgency purposes (see Table 2.5). However, for military aid, the inclusion of a variable accounting for the Reagan presidency yields a negative result for the amount of military aid given in Latin America and in the rest of the world, but the probability is only significant for Latin America. In sum, internal threats are only associated with a larger amount of military aid in Latin America, but not with economic aid.

Accounting for sample selection bias. Since we show in the previous paragraphs that the fact of receiving aid is not random, there is a potential for selection bias. To account for this we also estimate a two step Heckman selection model (Heckman, 1979). In Table 2.6 we include the main results that include both the selection equation ($Pr(aid)$) and the outcome equation ($lnEaid$).

It is worth noting that we are not relying solely upon the non linearity in the functional form of the selection equation. Columns (2) and (4) in Table 2.6 include the exclusion restrictions ($USworld$ and $USresources$) that provide a more robust

Table 2.5: **Guerrilla Effect on Aid Provision to Latin America and the World: 1946-2001**

	$lnEconomicAid$ (1)	$Pr(Aid)$ (2)
US intervention (Dummy)	1.142***	0.061
	(0.242)	(0.046)
$USintervention * LAC$	-1.346***	-0.078
	(0.326)	(0.064)
Political Regime (W) (Index 0-1 Democratic)	0.332	-0.071
	(1.476)	(0.217)
$PoliticalRegime * LAC$	-9.144***	-0.316
	(2.639)	(0.490)
Ideology (Index 0-10 conservative)	0.378	-0.017
	(0.610)	(0.031)
$Ideology * LAC$	-0.141	-0.070
	(0.345)	(0.067)
$W * Ideo$	0.057	0.006
	(0.257)	(0.038)
$W * Ideo * LAC$	1.630***	0.055
	(0.461)	(0.085)
$W * Intervention$	-0.215	0.308***
	(0.462)	(0.076)
$W * Intervention * LAC$	0.848	-0.178
	(0.619)	(0.114)
Guerrilla	0.015	-0.007
	(0.039)	(0.007)
$Guerrilla * LAC$	0.078	0.008
	(0.060)	(0.012)
Reagan	0.455	0.042*
	(0.484)	(0.022)
$Reagan * LAC$	-1.475***	-0.006
	(0.230)	(0.043)
Selectorate size (eS) (Index 0-1 Influence)	0.030	-0.003
	(0.121)	(0.024)
Share of government GDP (RB)	4.312***	0.085
	(0.409)	(0.075)
RB^2	-0.200***	-0.008***
	(0.016)	(0.003)
Country fixed effects	Yes	Yes
Year fixed effects	Yes	Yes
Observations	2,094	2,957
R-squared	0.204	0.054

Controls not shown: Ln pop, alignment with US (tau), Soviet Influence, amount of trade, new leader and years in office.

Notes: Robust standard errors clustered at the country level are shown in parentheses. *** is significant at the 1% level; ** is significant at the 5% level; * is significant at the 10% level.

identification. According to Bueno de Mesquita and Smith (2007), the characteristics of the donor country (US) represent valid exclusion restrictions that allows us to identify the Heckman model. Our findings in Table 2.6 are consistent with our main results, which suggest that for Latin America, ideology is an important factor in determining the amount of aid and the probability of receiving economic aid in the region. The coefficients show the same signs and are still significant, which validates our previous results.

Table 2.6: **Heckman Model of the Propensity of the US to Provide Aid to Latin America and the World, 1960-2001**

	Outcome Equation		**Selection Equation**	
	World	Latin America	World	Latin America
	(1)	(2)	(3)	(4)
US intervention (Dummy)	1.091***	-0.383	1.520**	-1.229
	(0.217)	(0.474)	(0.649)	(1.868)
Political Regime (W) (Index 0-1 Democratic)	0.559	-12.065***	-2.217	-40.226
	(1.310)	(4.688)	(2.986)	(12.297)
$W * Ideo$	0.011	2.284***	0.408	6.706
	(0.228)	(0.817)	(0.523)	(2.049)
$W * Intervention$	0.003	0.252	-0.820	4.881
	(0.412)	(0.869)	(0.836)	(3.569)
Selectorate size (eS) (Index 0-1 Influence)	0.009	-0.209	-0.431	5.561
	(0.117)	(0.504)	(0.398)	(2.293)
Share of government GDP (RB)	2.960***	7.466***	-4.344**	-16.366
	(0.416)	(2.699)	(1.381)	(8.217)
RB^2	-0.146***	-0.309***	0.124*	0.687
	(0.017)	(0.108)	(0.051)	(0.350)
$USworld$			36.915	-837.767
			(76.748)	(114.021)
$USresources$			3.482	-148.836
			(13.749)	(3.960)
Observations	2,290	669	2,290	669

Controls not shown: Ln pop, alignment with US (tau), Soviet Influence, amount of trade.
Notes: Robust standard errors clustered at the country level are shown in parentheses. *** is significant at the 1% level; ** is significant at the 5% level; * is significant at the 10% level.

2.6 Conclusions

This paper has examined how different factors affect US aid provision to Latin America and the rest of the World using a detailed longitudinal cross-country analysis. We present evidence showing that US foreign aid to Latin America follows a different pattern than the one observed for the rest of the world. Yet, contrary

to previous studies, we find that the amount of foreign aid allocation to the region is importantly influenced by the ideology of the United States government rather than by recipient characteristics such as institutions or internal threats (e.g. guerrillas). These results are robust to controlling for a variety of alternative hypotheses, including economic controls as well as differential trends across countries.

We also present evidence on how foreign aid allocations are distributed to other parts of the world in order to appropriately compare the patterns observed. After controlling for numerous political, economic and geographical factors, we find that the US did not always use aid as a transfer to autocracies in exchange for policy concessions in Latin America. With respect to the rest of the world, aid appears to be directed to more democratic countries. Such comparison is important in the case of regional studies since it allows us to not assume any differential policy, but rather to show with evidence that such differences exist

We find that donor characteristics influence the extent to which foreign aid was used to obtain policy concessions from Latin America during 1946-2001. Moreover, the ideology of the US government does not explain much of US aid provision to the rest of world but it does so to Latin America. In addition, aid provision to the rest of the world is affected by CIA interventions more than by the type of political regime in the receiving country, while CIA interventions do not affect the amount of aid given to Latin American countries.

This paper also calls for a more nuanced view on commonly held beliefs about US foreign policy towards Latin America. For instance, the alleged militarization of the region due to US military aid or the use of foreign aid for counterinsurgency efforts appears unsubstantiated. Rather, US - Latin America relations have seen that aid is targeted to more democratic institutions under conservative governments. Therefore, at least some part of the effect observed can be attributed to the preferences, goals and objectives of US administrations which vary according to the economic or ideological view of those in power.

Appendices

2.A Results for Military Aid

Table 2.A.1: **Military Aid Provision to Latin America and the World: 1946-2001**

	$lnMilitaryAid$		$Pr(MilitaryAid)$	
	(1)	(2)	(3)	(4)
US intervention (Dummy)		0.292		0.312***
		(0.290)		(0.038)
$USintervention * LAC$		0.430		-0.162***
		(0.344)		(0.052)
Political Regime (W) (Index 0-1 Democratic)	0.752***	0.754**	-0.060	0.026
	(0.258)	(0.295)	(0.042)	(0.046)
$PoliticalRegime * LAC$	-0.885**	-0.560	0.150**	0.206***
	(0.375)	(0.407)	(0.066)	(0.070)
Selectorate size (eS) (Index 0-1 Influence)	-0.006	-0.099	0.063**	0.050
	(0.176)	(0.183)	(0.029)	(0.030)
Share of government GDP (RB)	4.313***	4.457***	0.994***	1.043***
	(0.669)	(0.728)	(0.081)	(0.086)
RB^2	-0.155***	-0.156***	-0.041***	-0.043***
	(0.024)	(0.026)	(0.003)	(0.003)
Country fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Observations	1,875	1,623	3,929	3,536
R-squared	0.161	0.164	0.233	0.248

Controls not shown: Ln pop, alignment with US (tau), Civil war, Soviet Influence, amount of trade, contiguous border, common language, GATT participant, regional trade agreement.
Notes: Robust standard errors clustered at the country level are shown in parentheses. *** is significant at the 1% level; ** is significant at the 5% level; * is significant at the 10% level.

Table 2.A.2: **US Ideology on Military Aid Provision to Latin America and the World: 1960-2001**

	$lnMilitaryAid$ (1)	$Pr(MilitaryAid)$ (2)
US intervention (Dummy)	1.480***	0.305***
	(0.424)	(0.059)
$USintervention * LAC$	-1.409***	-0.283***
	(0.530)	(0.084)
Political Regime (W) (Index 0-1 Democratic)	1.038	1.684***
	(2.579)	(0.291)
$PoliticalRegime * LAC$	-0.044	-4.256***
	(4.388)	(0.639)
Ideology (Index 0-10 Conservative)	-1.731***	0.074
	(0.580)	(0.076)
$Ideology * LAC$	-0.213	-0.679***
	(0.496)	(0.072)
$W * Ideo$	0.071	-0.292***
	(0.443)	(0.051)
$W * Ideo * LAC$	-0.331	0.769***
	(0.772)	(0.111)
$W * Intervention$	-1.931***	-0.009
	(0.733)	(0.099)
$W * Intervention * LAC$	3.042***	0.185
	(0.913)	(0.148)
Selectorate size (eS) (Index 0-1 Influence)	-0.078	0.034
	(0.193)	(0.030)
Share of government GDP (RB)	4.178***	1.120***
	(0.797)	(0.095)
RB^2	-0.151***	-0.048***
	(0.029)	(0.004)
Country fixed effects	Yes	Yes
Year fixed effects	Yes	Yes
Observations	1,392	3,072
R-squared	0.148	0.252

Controls not shown: Ln pop, alignment with US (tau), Civil war, Soviet Influence, amount of trade, contiguous border, common language, GATT participant, regional trade agreement.

Notes: Robust standard errors clustered at the country level are shown in parentheses. *** is significant at the 1% level; ** is significant at the 5% level; * is significant at the 10% level.

Table 2.A.3: **Guerrilla Effect on Military Aid Provision to Latin America and the World: 1946-2001**

	$lnMilitaryAid$ (1)	$Pr(MilitaryAid)$ (2)
US intervention (Dummy)	1.565***	0.321***
	(0.425)	(0.060)
$USintervention * LAC$	-1.484***	-0.301***
	(0.530)	(0.084)
Political Regime (W) (Index 0-1 Democratic)	0.866	1.729***
	(2.579)	(0.293)
$PoliticalRegime * LAC$	0.350	-4.382***
	(4.433)	(0.641)
Ideology (Index 0-10 Conservative)	-1.649***	-0.446***
	(0.583)	(0.124)
$Ideology * LAC$	0.704	-0.471***
	(0.574)	(0.087)
$W * Ideo$	0.099	-0.294***
	(0.443)	(0.051)
$W * Ideo * LAC$	-0.365	0.792***
	(0.780)	(0.112)
$W * Intervention$	-1.959***	-0.017
	(0.733)	(0.099)
$W * Intervention * LAC$	2.892***	0.136
	(0.915)	(0.148)
Guerrilla	-0.088*	0.021**
	(0.047)	(0.009)
$Guerrilla * LAC$	0.084	0.010
	(0.076)	(0.016)
$Reagan$	-3.446***	-0.093
	(0.711)	(0.092)
$Reagan * LAC$	-1.094***	-0.248***
	(0.369)	(0.055)
Selectorate size (eS) (Index 0-1 Influence)	-0.092	0.029
	(0.194)	(0.030)
Share of government GDP (RB)	3.955***	1.052***
	(0.807)	(0.096)
RB^2	-0.147***	-0.045***
	(0.029)	(0.004)
Country fixed effects	Yes	Yes
Year fixed effects	Yes	Yes
Observations	1,381	3,019
R-squared	0.158	0.258

Controls not shown: Ln pop, alignment with US (tau), Soviet Influence, amount of trade, contiguous border, common language, GATT participant, regional trade agreement.

Notes: Robust standard errors clustered at the country level are shown in parentheses. *** is significant at the 1% level; ** is significant at the 5% level; * is significant at the 10% level.

Table 2.A.4: **US Ideology and Aid Provision**

	lnEconomicAid - no Exports		*lnEconomicAid* - with Exports	
	World	Latin America	World	Latin America
	(1)	(2)	(3)	(4)
US intervention (Dummy)	0.862**	-0.121	0.667*	-0.122
	(0.367)	(0.324)	(0.365)	(0.327)
Political Regime (W) (Index 0-1 Democratic)	0.172	-4.275**	0.129	-4.258**
	(1.485)	(1.890)	(1.459)	(1.873)
$W * Ideo$	0.124	0.864***	0.099	0.858***
	(0.239)	(0.312)	(0.237)	(0.309)
$W * Intervention$	-0.586	0.449	-0.286	0.461
	(1.001)	(0.496)	(0.995)	(0.498)
Selectorate size (eS) (Index 0-1 Influence)	0.147	-0.173	0.202	-0.179
	(0.198)	(0.313)	(0.188)	(0.317)
US Exports			0.371***	0.170
			(0.089)	(0.240)
Share of government GDP (RB)	3.237***	12.353***	3.229***	12.329***
	(0.788)	(2.310)	(0.728)	(2.317)
RB^2	-0.157***	-0.502***	-0.162***	-0.502***
	(0.033)	(0.088)	(0.030)	(0.088)
Country fixed effects	Yes	Yes	Yes	Yes
Observations	1,494	604	1,494	604

Controls not shown: Ln pop, alignment with US (tau), Soviet Influence, amount of trade, contiguous border, common language, GATT participant, regional trade agreement, new leader and years in office.

Notes: Newey-West robust standard errors in parentheses (30 lags). *** is significant at the 1% level; ** is significant at the 5% level; * is significant at the 10% level.

2.B Classification of countries by region

Table 2.B.1: **Classification: World Bank**

Region	Countries
East Asia and Pacific	Samoa, Cambodia, China, Taiwan, Fiji, Indonesia, South Korea, North Korea, Laos, Malaysia, Marshall Islands, Myanmar, Palau, Papua New Guinea, Philippines, Solomon Islands, Thailand, Vietnam
Middle East and North Africa	Algeria, Djibouti, Egypt, Iran, Iraq, Israel, Jordan, Lebanon, Libya, Morocco, Syria, Tunisia, Yemen.
Sub-Saharan Africa	Angola, Benin, Botswana, Burkina Faso, Burundi, Benin, Botswana, Cameroon, Guinea, Gambia, Ghana, Guinea Bissau, Rwanda, Senegal, Ethiopia, Gabon, Seychelles, Zambia, Zimbabwe, Nigeria, Central African Republic, Chad, Comoro Islands, Congo DR, Congo REP, Cote D'Ivoire, Cape Verde, Eritrea, Kenya, Lesotho, Liberia, Madagascar, Malawi, Mali, Mauritania, Mauritius, Mozambique, Namibia, Niger, Sierra Leone, South Africa, Somalia, Sudan, Swaziland, Sao Tome & Principe, Tanzania, Togo, Uganda.
Latin America and the Caribbean	Antigua, Argentina, Belize, Brazil, Chile, Colombia, Costa Rica, Jamaica, Dominican Republic, Dominica, Ecuador, El Salvador, Grenada, Guatemala, Guyana, Haiti, Honduras, Mexico, Nicaragua, Panama, Paraguay, Peru, St. Kitts/Nevis, St. Vincent, Suriname, Uruguay, Venezuela.

Bibliography

Alesina, Alberto and David Dollar (2000). "Who gives foreign aid to whom and why?" In: *Journal of Economic Growth* 5, pp. 33–63.

Berger, Daniel et al. (2013). "Commercial Imperialism? Political Influence and Trade During the Cold War". In: *American Economic Review* 103.2, pp. 863–896.

Breuning, Marijke (1995). "Words and deeds: Foreign assistance rhetoric and policy behavior in the Netherlands, Belgium, and the United Kingdom". In: *International Studies Quarterly* 39, pp. 235–254.

Bueno de Mesquita, Bruce (1981). *The war trap*. New Haven : Yale University Press.

Bueno de Mesquita, Bruce and Alastair Smith (2007). "Foreign Aid and Policy Concessions". In: *Journal of Conflict Resolution* 51, pp. 251–284.

— (2009). "A political economy of aid". In: *International Organization* 63.2, pp. 309–340.

Cingranelli, David L. and Thomas E. Pasquarello (1985). "Human Rights Practices and the Distribution of US Foreign Aid to Latin American Countries". In: *American Journal of Political Science* 29.3, pp. 539–563.

Dunning, Thad (2004). "Conditioning the effects of aid: Cold war politics, donor credibility, and democracy in Africa". In: *International Organization* 58.2, pp. 409–23.

Fleck, Robert K. and Christopher Kilby (2010). "Changing aid regimes? US foreign aid from the cold war to the war on terror". In: *Journal of Development Economics* 91.2, pp. 185–197.

Goldstein, Markus and Todd Moss (2005). "Compassionate conservatives or conservative compassionates? US political parties and bilateral foreign assistance to Africa". In: *Journal of Development Studies* 41.7, pp. 1288–1302.

Heckman, J. (1979). "Sample selection bias as a specification error". In: *Econometrica* 47.1, pp. 153–61.

Lagon, Mark P. (1992). "The International System and the Reagan Doctrine: Can Realism Explain Aid to 'Freedom Fighters'?" In: *British Journal of Political Science* 22.1, pp. 39–70.

Lumsdaine, David H. (1993). "Moral vision in international politics: The foreign aid regime 1949-1989". In: *Princeton, N.J.: Princeton University Press.*

Maoz, Zeev and Bruce Russett (1993). "Normative and Structural Causes of Democratic Peace, 1946-1986". In: *Americal Political Science Review* 87, pp. 624–638.

McCormick, James and Neil Mitchell (1988). "Is US aid really linked to human rights in Latin America?" In: *American Journal of Political Science* 32.1, pp. 231–239.

Meernik, James, Eric L. Krueger, and Steven C. Poe (1998). "Testing models of US foreign policy during and after the cold war". In: *The Journal of Politics* 60.1, pp. 63–85.

Milner, H. and D. Tingley (2010). "The Domestic Politics of Foreign Aid: American Legislators and the Politics of Donor Countries". In: *Economics and Politics* 22.2, pp. 200–232.

Muller, E.N. (1985). "Dependent Economic Development, Aid Dependence on the United States, and Democratic Breakdown in the Third World". In: *International Studies Quarterly* 29.4, pp. 445–469.

Packenham, Robert A. (1973). *Liberal America and the Third World: Political development ideas in foreign aid and social science.* Princeton, N.J., Princeton University Press.

Sanchez, Peter (2003). "Bringing the International Back in: US Hegemonic Maintenance and Latin America's Democratic Breakdown in the 1960s and 1970s". In: *International Politics* 40, pp. 223–247.

Tingley, Dustin (2010). "Donors and domestic politics: Political Influences on foreign aid effort". In: *Quarterly Review of Economics and Finance* 50.1, pp. 40–49.

Chapter 3

Corporate Social Responsibility and the Economy

Abstract

What are the determinants of a firm to engage in Corporate Social Responsibility (CSR) and under what conditions is it likely to occur? This paper presents evidence of two possible mechanisms that affect the decision of a firm to engage in CSR: the role of growth in value added within an industry and workers' attitudes. The results suggest that firms engage in CSR practices in times of economic prosperity. I also provide further evidence that peer effects are important on the decision of a firm to participate in CSR. When the proportion of firms engaged in CSR increases, the probability of a firm that hasn't engaged before increases. With respect to workers' attitudes, this paper provides evidence of a weak link between labour force preferences and a firm's decision to engage in CSR.

3.1 Introduction

During the last few decades, there has been an increasing interest by firms, governments and society to include Corporate Social Responsibility (CSR) into the firm's business strategy (e.g. Siegel and Vitaliano 2007; Lankoski 2009; Lundgren 2011; Jenkins 2005; Orlitzky, Schmidt, and Rynes 2003, among others). Such interest in corporate social responsibility is considered the "latest manifestation of a longstanding debate over the relationship between business and society" (Jenkins, 2005). The recent attention by the media and academic literature has mostly focused on the environmental impact of firms. Yet, this notion has expanded to include issues such as human rights, community strengthening, sexual diversity, employees' rights, among others. As Kitzmueller and Shimshack (2012) recognize, "Corporate Social Responsibility is an important economic phenomenon with extensive implications for firms, employees, consumers, investors, governments and NGOs alike." This paper provides empirical evidence that helps to understand under what conditions corporate social responsibility could take place.

What are the determinants of a firm to engage in CSR and under what conditions is it likely to occur? The response of CSR activities to business cycle fluctuations is key to understanding how firms behave. Cyclical fluctuations in CSR activities may help explain the way the prevailing macroeconomic context affects firm's decisions on CSR (Branca, Pina, and Catalao-Lopes, 2012). Unfortunately, there hasn't been enough research about the actual fluctuations in CSR investments. There is a small empirical literature discussing the determinants of CSR both through the institutional side and strategic side, but the lessons from this literature are mixed. For the purpose of this paper, I consider the definition of CSR to be "actions that, to some degree, imply corporate beyond-compliance behavior in the social and/or the environmental arena" (Lundgren, 2011). In other words, CSR implies a behavior beyond the prevailing legal or regulatory requirements. As noted, this includes not only environmental measures, but also community building,

human rights protection, employee relations, gender representation and diversity, among others.

In this paper, I take a closer look at the cyclicality of CSR activities and other determinants to CSR that have not been accounted for in previous literature. To do so, I use COMPUSTAT and KLD to construct a firm-level panel dataset containing CSR activities and firm characteristics for a period spanning 12 years (1998-2010). I complement COMPUSTAT with data on industry value added using the US Department of Commerce (Bureau of Economic Analysis) data. Finally, for the last part of the paper I include data on labor force attitudes towards key areas of CSR (e.g. environment, human rights and diversity) using American National Election Surveys (ANES) to assess the linkages (or lack thereof) between workers and firms.

This paper makes a contribution to the broader literature by narrowing the gap between the evidence available and the information required for making theoretical assumptions about the cyclicality of CSR activities. In particular, the paper exploits the information contained in COMPUSTAT and ANES to asses two possible mechanisms that affect the decision of a firm to engage in CSR: the role of the business cycle and workers' attitudes. By analyzing these two mechanisms I do not intend to make an overly deterministic claim that these are the only two mechanisms responsible for the adoption of social responsibility, but I find that the macroeconomic conditions do affect the entrepreneurial decisions to engage in CSR, and workers' attitudes can drive the firm's decisions to implement CSR activities, particularly of diversity CSR practices within a firm.

The results of the paper indicate that overall there is a strong and positive relationship between the growth in value added in the firm's industry and the decision to engage in CSR. This relationship, while strong for most of the firms, does not hold for top-performing firms, who seem to engage in CSR regardless of their industry value added. Furthermore, peer effects are relevant for a firm's

decision to engage in CSR up to a point when the agglomeration of firms engaging in CSR is no longer relevant for another firm to engage in CSR within the same industry.

Finally, the results also indicate that overall workers preferences are not driving a firm to engage in CSR, proving a weak link between firm's engagement in CSR and workers' attitudes. However, there is evidence that attitudes towards LGBT have enough relevance to influence the decision of a firm to engage in diversity CSR activities.

Relation to the literature To my knowledge, there are very few papers which estimate the effect of macroeconomic conditions and workers' attitudes on CSR adoption. For example, while Branca, Pina, and Catalao-Lopes (2012) show that firms are expected to reduce CSR spending when the business cycle is unfavorable. They only analyze the corporate giving dimension of CSR without taking into consideration other CSR activities such as environmental, human rights, diversity, etc. Campbell (2007) argues that firms will not behave socially responsible when they are operating in an unhealthy economic environment, however he does not provide empirical evidence of this statement. Hsiang-Lin, Hsiang-Hsuan, and Tzu-Yin (2010) empirically test Campbell's argument by including the inflation rate, industrial production index and the consumer confidence index as the economic variables. They find that firms located in countries with higher consumer confidence levels and lower inflationary levels will act in more socially responsible ways, but they only estimate the results for financial firms in a country level setting.

The closest paper to mine in terms of the cyclicality of CSR is Branca, Pina, and Catalao-Lopes (2012). Although we look at related issues in a similar context, we have very different approaches. Methodologically, Branca, Pina, and Catalao-Lopes (2012) take a times series approach by estimating the *cross correlation pattern* of CSR giving and firm revenues with real GDP. They find a significant co-movement

between real GDP and CSR giving. In contrast, I use panel data to estimate the relationship between the growth in industry value added and the probability of a firm to engage in CSR activities.

My approach has several advantages and provides new insights on the effects of macroeconomic conditions on the decision of companies' to act in socially responsible ways. First, I offer more precise and reliable estimates on the impact of the business cycle on CSR by using firm's data. By using IRS data, Branca, Pina, and Catalao-Lopes (2012) focus solely on giving as the only measure of CSR. But data limitations do not allow her to control for any other firm's characteristics.

From the workers' preferences perspective, the closest paper is Marquis, Glynn, and Davis (2007). They provide a framework to look at how the community and culture pressure firms to implement CSR. Marquis, Glynn, and Davis (2007) suggest that communities are important influencers on corporate action. There are two main reasons for this argument. First, the set of rules and norms that are in place in the community serve also as a benchmark for legitimizing CSR (see also Marquis, Glynn, and Davis (2007); Joshua D Margolis and James P Walsh (2003)). Second, CSR activities are commonly oriented towards the community where the corporation's executive reside (Marquis, Glynn, and Davis, 2007). However, Marquis, Glynn, and Davis (2007) does not provide any empirical analysis of their framework. Several studies in the US have shown that this is the case by using small samples of firms across the country. Most of these research are study cases, for example Galaskiewicz and Wasserman (1989) looks at local giving patterns of Minneapolis based companies and how they are influenced by network effects.

To the best of my knowledge, there is no study that includes workers' attitudes from ANES into the analysis of Corporate Social Responsibility and data on firms from COMPUSTAT and KLD for the period that I consider (1998-2010).

The remainder of the paper is organized as follows. Section 2 provides a broader literature review on the determinants of CSR, Section 3 includes a description of

the dataset and some basic facts. The empirical methodology and the main results of the paper are described in Section 4 and 5 accordingly. Section 6 concludes.

3.2 Literature Review

Research on CSR has mainly focused on investigating the relationship between corporate social responsibility and corporate financial performance (Joshua D Margolis and James P Walsh (2003), Bagnoli and Watts (2003),Orlitzky, Schmidt, and Rynes (2003), among others). Most of these studies focus on determining the extent to which socially responsible corporate behavior affects financial performance. These studies have established a positive relationship between firm's economic performance and social responsibility by attracting socially responsible consumers (Bagnoli and Watts, 2003), reducing the threat of regulation (Maxwell, Thomas P. Lyon, and Hackett, 1998), improving their reputation, and reducing concern from non governmental organizations (see Joshua D. Margolis, Elfenbein, and James P. Walsh (2007), McWilliams, Siegel, and Wright (2006), McWilliams and Siegel (2001), Hull and Rothenberg (2008), Lundgren (2011), Reinhardt, Stavins, and Vietor (2008), Baron (2001), Thomas Peyton Lyon and Maxwell (2007)).

Joshua D Margolis and James P Walsh (2003) review this literature and find that out of 127 published studies that empirically analyze the relationship between firms' corporate social responsibility and their financial performance, 85 percent of the studies treat CSR as independent variable, and only 22 studies treat CSR as dependent variable. Moreover, they show that there has been a tendency to ignore several factors, other than corporate financial performance, which may affect CSR. This leaves room to analyze factors other than corporate financial performance that may affect corporate social responsibility.

As noted above, most of the literature on corporate social responsibility does not explore whether macroeconomic conditions affect the likelihood of firms to

behave in socially responsible ways. The increasing number of empirical studies attempt to establish models of strategic CSR. For example McWilliams and Siegel (2001) employ regression analysis to evaluate the relationship between profitability and CSR using the overall firm-level index of CSR data from KLD. Most of the research that empirically test this relationship uses CSR as an explanatory variable, including a few other characteristics of the firm that may affect the firm's profitability measures. Among this research, Baron (2001) and McWilliams and Siegel (2001) were the first to explicitly model "profit maximizing" CSR.

Most of the literature follows the same line, although there are a few that include other characteristics. For example Navarro (1988) looks at the effect of the tax structure on corporate giving; Marquis, Glynn, and Davis (2007) show that pressures at the community level shape corporate action; Bagnoli and Watts (2003) argues that the provision of a good depends on consumers' willingness to pay and on the structure of competition.

Macroeconomic conditions and Corporate Social Responsibility. Among the few papers that look at the relationship between the macroeconomic environment and CSR are: Branca, Pina, and Catalao-Lopes (2012) that test empirically the relationship between corporate giving and the business cycle. Their findings suggest that with adverse economic conditions, firms are expected to reduce CSR spending, independently of the market structure. Ullmann (1985) states that "in periods of low profitability, economic demands will have priority over social demands..."; and Campbell (2007) argues that corporations will be less likely to behave in a socially responsible way when the operating environment is limited.

While most of these studies are only normative, Branca, Pina, and Catalao-Lopes (2012) empirically tests the relationship between firm's giving and the macroeconomic environment. Similarly, I examine the role of macroeconomic conditions on the firm's decision to engage in CSR.

Worker's Attitudes and Corporate Social Responsibility. For worker's attitudes, the role of individual or community attitudes towards CSR is less explored. There is one study from Marquis, Glynn, and Davis (2007) that explains how the community in which firms are embedded affects the nature and level of corporate social actions. They propose that community-level cultural forces affect the nature of corporate social action, but they do not empirically test their hypothesis. Kitzmueller and Shimshack (2012), assert that the labor force social preferences may translate into demand for corporate social responsibility.

For community pressure I use data from US national surveys[1] to analyze the possibility that workers' pressure toward firms may have a larger role for adopting CSR measures –with respect to specific CSR activities, such as environment, diversity, employment, among others–.

3.3 Data

Data for this research consists of a panel dataset of 4,785 large publicly traded companies listed on the main international stock exchanges, with observations from 1998 to 2010. I have four main sources of data to look at Corporate Social Responsibility; it's relationship with the macroeconomic environment and to labor force preferences. First, KLD from 1998 to 2010 (12 years). The second is COMPUSTAT, which covers the same years. The third data source is industry value added from the US Department of Commerce (Bureau of Economic Analysis) with observations from 1998 to 2010. Finally, I include data on workers opinions towards key areas of CSR (e.g. environment, human rights and diversity) using American National Election Surveys (ANES) from 1998 to 2009.

KLD (Kinder, Lydenberg, Domini Research and Analytics). Is an annual dataset on environmental, social, and governance ratings. It is published at the end of each calendar year. It identifies the company information (Name,

[1]Data from national surveys provided by the American National Elections Survey

Ticker, unique security identifiers), and provides 50 CSR indicators organized in seven categories: the environment, community, human rights, employees, diversity, production and corporate governance. Table 3.1 describes the seven areas covered by the KLD and their components.[2] The number of firms in the sample has a substantial increase in 2003, and from that year onward KLD added CSR ratings for all firms belonging to the $Russell1000Index$.

COMPUSTAT. Is an annual database of financial, statistical, and market information on global companies. The characteristics include indicators such as total assets, earnings, industry classification, sales, number of employees, among others. Such characteristics would most likely account for time varying conditions faced by the firm.

Value Added by Industry. This data comes form the Bureau of Economic Analysis. It includes the value added by industry sector. The data is available from 1998 to 2010.

Labour Force Opinion. It comes from the American National Election Studies survey which conducts a bi-annual survey on the state of individual's attitudes towards an array of issues. In particular, I use information about workers' attitudes (on a 0 to 100 scale, 100 being positive) towards LGBT rights, unions and environmentalists. Such data is coded by state and matched to the state in which the firm is based. Although this might not directly capture the workers' opinions each individual firm faces, it would provide a proxy for the type of attitudes prevailing in the American electorate as a whole.

The purpose of the paper is to analyze the effect of economic activity on the decision to engage in CSR activities. In addition to economic activity (measured here with the growth of value added), there are other factors that can influence CSR activities. These factors should also be included in the regression to control for their potential influence on CSR decisions. According to Belu and Manescu

[2]For more detailed explanation visit www.msci.com.

Table 3.1: **CSR Indicators and Components**

Indicator	**Components**
Environmental	Beneficial Products and Services
	Pollution Prevention
	Recycling
	Clean Energy
	Management Systems
	Other
Community	Charitable Giving
	Innovative Giving
	Community Engagement
	Other
Human Rights	Indigenous Peoples Relations Strength
	Human Rights Policies and initiative
Employee Relations	Union Relations
	Cash Profit Sharing
	Employee Involvement
	Health and Safety
	Supply Chain Policies, Programs and Initiative
	Other
Diversity	Representation (women and minorities)
	Board of Directors (women and minorities)
	Work/Life Benefits
	Women and Minority Contracting
	Gay and Lesbian Policies
	Employment of Under represented Groups
	Other
Product	Quality
	Benefits to Economically Disadvantaged
	Access to Capital
	Product Safety
	Marketing/Contracting Concern
	Antitrust
	Customer Relations Concern
Governance	Reporting Quality
	Public Policy

Source: www.msci.com

(2013) firm size is relevant for economic performance. The bigger the firm, higher variation on profitability of the firms. Market structure is important for CSR, since it affects the decision of adopting CSR (Bagnoli and Watts, 2003). Results can vary significantly depending on the market structure. If markets are competitive, monopolistic or oligopolistic, market power enables some firms to always to earn unusual returns.

3.4 Methodology

In the empirical implementation of this paper I run a set of fixed effects regressions to estimate the relation between CSR adoption and the business cycle as well as the relation between CSR and labor force opinions. All results presented account for heteroskedasticity and serial correlation. I assume a *linear probability model* (LPM), where changes in the explanatory variables have a linear effect on the change in the probability of adopting CSR and no arbitrary choice of a nonlinear regression function is made (Wooldridge, 2002). Econometrically, it is difficult to estimate using alternatives like fixed effects Probit using standard methodologies.[3] Hence, I estimate equation 3.1 with a linear probability model in which: CSR_{ijt} is an indicator variable that equals one when firm i from industry j engages in any CSR activity at time t, and zero otherwise; $Growth_{jt-1}$ is the growth in value added observed for industry j at time $t-1$; X_{ijt} is a set of firm characteristics (see below); t is a linear trend capturing the overall growing importance of CSR over time; and u_{ijt} is an i.i.d idiosyncratic error term.

While T is not sufficiently large in our sample, I also allow u_{ijt} to follow an AR(1) process: $u_{ij,t} = \rho u_{ij,t-1} + \epsilon_{ijt}$.[4] The magnitude of the estimated coefficient

[3] A Logit model would provide adequate standard errors but with a biased estimate, since I wouldn't be accounting for unobserved heterogeneity. I decided to estimate a fixed effects model to have an unbiased estimate.

[4] I also allow for an AR(2) process and heteroskedasticity in our unbalanced panel using xtAR-GLS and results do not differ from the ones assuming the error term follows an AR(1) process (column (3) in Table 3.3). Furthermore I also present Driscoll-Kraay standard errors (Driscoll

of interest (β) does not change dramatically and remains statistically significant.

$$CSR_{ijt} = \alpha_i + \beta Growth_{jt-1} + \delta X_{ijt} + \gamma t + u_{ijt} \tag{3.1}$$

Peer effects. To analyze how the proportion of firms within the same industry affect the decision to engage in CSR for a particular firm in the same industry, I estimate equation 3.2, where α_i and X_{ijt} are defined as before, I further include year fixed effects (γ), and interest lies in θ and Θ.

$$CSR_{ijt} = \alpha_i + \gamma_t + \theta pfirms + \Theta pfirms^2 + \delta X_{ijt} + u_{ijt} \tag{3.2}$$

Labor Force opinions. I exploit state-level variation in workers attitudes (towards specific issues) to assess if these attitudes have an effect on the likelihood of firms to engage in (overall and thematic) CSR. I estimate equation 3.3 using a fixed effects LPM.

$$CSR_{ist} = \alpha_i + \delta X_{ist} + \eta LGBT_{st} + \lambda Unions_{st} + \rho Env_{st} + \gamma_t + u_{ist} \tag{3.3}$$

Controls. I consider a number of control variables (X_{ijt}) that are related to the firm's characteristics, and are always included in the specialized literature related to CSR. They include firm size; return on assets interacted by number of employees, as another measure of firm's size, advertising intensity, R&D expenses, and the Hirschman-Herfindahl Index (HHI).

Return on Assets (ROA) is a profitability measure that expresses the amount of net income plus (after tax) interest payments but before preferred dividend per unit of average current and last year's assets (Belu and Manescu, 2013). Firm size is measured as the number of employees of the firm; advertising intensity is defined as the log advertising expenses; R&D intensity is the log of R&D expenses. The

and Kraay, 1998) that are robust to both, *spatial* and *serial* autocorrelation of $u_{ij,t}$

Hirschman-Herfindahl Index for each industry is defined at the six-digit NAICS code level. It is computed by adding the square of the market share of all players operating in an industry a given year.[5] As displayed in Table 3.2, the control variables have values similar to the existing literature.

Summary Statistics. In total, I start with 4,785 firms, resulting in an unbalanced panel with an average length of 5.5 years. An important feature of the empirical strategy is the inclusion of firm fixed effects. To have enough within-firm variation, I further refine the sample to firms with at least 4 years of data. This results in a final sample of 3,068 firms with an average length of 6 years. Table 3.2 includes summary statistics for the sample.

Table 3.2: **Summary Statistics**

Variable	Mean	Std. Dev.	N
ROA (Return on Assets in logs)	0.771	0.18	15,865
Emp (Number of employees)	16.355	56.96	23,063
$ROA \times Emp$	17.219	57.916	15,743
HHI (Herfindahl Index)	0.187	0.177	23,368
R&D (log of R&D expenditures)	0.072	0.527	23,368
Advert (log of Advertisement expenditures)	4.113	7.275	23,368
$Growth_{jt-1}$ (Growth in Value added)	0.03	0.111	19,809
LGBT (CSR diversity index)	0.5	0.07	14,442
Unions (Union relations index)	0.564	0.055	14,442
Environment (CSR environment index)	0.65	0.046	13,601
$Pfirms$ (# firms CSR/#firms industry)	0.560	0.172	23,368

3.5 Results

3.5.1 CSR and Growth by Industry.

In this section, I asses the relationship between the growth of value added and the probability of a firm engaging in CSR. My first estimation results are reported in Table 3.3. Column 1 reports the results of a linear probability model with fixed

[5]See Fernandez-Kranz and Santalo (2010), Belu and Manescu (2013), Liu and Wu, 2015, and Fernandez-Kranz and Santalo (2010) for further details.

effects and in the next columns I account for the possibility of misspecification due to autocorrelation in my data.

I find that the coefficient on the growth of value added, $Growth_{jt-1}$, is positive and statistically significant even after accounting for autocorrelation. Column 2 reports the estimates of equation 3.1 with autocorrelation of order AR1. I find that the coefficient on the growth of value added, $Growth_{jt-1}$, is positive and statistically significant. Also, given that the confidence intervals of the coefficients on $Growth_{jt-1}$ overlap, I will keep the most parsimonious model with no autocorrelation for the rest of the paper. For a further check of the results in Table 3.3, I include the difference of the dependent variable by looking at firm's transitions to engage in CSR or stop doing CSR activities; the results hold with positive and significant coefficients as seen in the Appendix Table 3.A.1.

Gazelles vs. Elephants. Even though the overall results reflect that the economic environment at the industrial level (i.e. $Growth_{jt-1}$) is important on the decision of a firm to engage in CSR, we estimate the heterogeneous treatment by the firm's economic performance.[6] I find that the $Growth_{jt-1}$ of the industry is insignificant for the best performers of the industry "Gazelles" (top 25 percent) to engage in CSR. For the "Elephants" (bottom 75 percent), it appears that even when controlling for firm's characteristics and a time trend, an increase in $Growth_{jt-1}$ leads to an increase in the probability to engage in CSR Campbell (2007).

Decision to transition to CSR ($D-CSR$). Table 3.B.3 shows the marginal effects of a multinomial logit where the probability to engage in CSR when there is a ten percentage increase in $Growth_{jt-1}$ the probability of adopting CSR increases by 1.7 percentage points (equivalent to 25 percent), with respect to not adopting. Similarly, when there is a ten percentage increase in $Growth_{jt-1}$ the probability of disengaging in CSR decreases by 4 percentage points (equivalent to 62 percent).

[6]Economic performance is measured as Return on Assets (ROA) in my sample.

Table 3.3: **Corporate Social Responsibility and Growth of Value Added by Industry**

Corporate Social Responsibility					
	(1) OLS	(2) Logit	(3)	(4) $AR(1)$	(5) $AR(\rho)$
$Growth_{jt-1}$	0.16***	1.53***	0.075***	0.10***	0.16***
	(0.057)	(0.31)	(0.029)	(0.029)	(0.035)
# Employees	-0.00080	0.0076	-0.00011	-0.00074	-0.00080**
	(0.00086)	(0.0084)	(0.0010)	(0.00047)	(0.00032)
$ROA \times Emp$	0.0011	0.0100	0.00048	0.0012***	0.0011***
	(0.00092)	(0.0084)	(0.0010)	(0.00040)	(0.00030)
HHI (Herfindahl Index)	-0.11	-0.94**	-0.13***	-0.19***	-0.11
	(0.092)	(0.39)	(0.050)	(0.059)	(0.071)
ln R&D expenditures	0.0084**	0.12	0.0060	0.0067*	0.0084
	(0.0038)	(0.079)	(0.0057)	(0.0034)	(0.0049)
ln Advertisement expenditures	0.0031**	0.031***	0.0049***	0.0037**	0.0031**
	(0.0016)	(0.012)	(0.0014)	(0.0015)	(0.0013)
Year	0.011***	0.12***		0.0070***	0.011**
	(0.0031)	(0.013)		(0.0013)	(0.0047)
Firm FE	YES	YES	YES	YES	
Observations	13,491	6,591	11,248	13,139	13,491
Number of id	2,243	1,003	2,144	2,123	
Number of groups					2,243

Notes: Robust standard errors clustered at the industry level are shown in parentheses. Variables not shown include firm fixed effects. *** is significant at the 1% level; ** is significant at the 5% level; * is significant at the 10% level.

Table 3.4: **Heterogeneous effects: Gazelles vs. Elephants**

Corporate Social Responsibility		
	(1) Elephants	(2) Gazelles
$Growth_{jt-1}$	0.23***	-0.018
	(0.063)	(0.071)
# Employees	-0.0017**	-0.00092
	(0.00083)	(0.0019)
$ROA \times Emp$	0.0019***	0.0010
	(0.00068)	(0.0021)
HHI (Herfindahl Index)	-0.14	-0.25*
	(0.11)	(0.14)
ln R&D expenditures	0.0075	0.011**
	(0.0057)	(0.0040)
ln Advertisement expenditures	0.0056***	-0.0079**
	(0.0017)	(0.0038)
Year	0.010***	0.0100**
	(0.0036)	(0.0039)
Firm FE	YES	YES
Observations	10,025	3,466
R-squared	0.016	0.020
Number of id	2,032	791

Notes: Robust standard errors clustered at the industry level are shown in parentheses. Variables not shown include firm fixed effects. *** is significant at the 1% level; ** is significant at the 5% level; * is significant at the 10% level.

Table 3.5: **Marginal effects: Multinomial logit**

Corporate Social Responsibility		
	(1) Engage	(2) Disengage
$Growth_{jt-1}$	0.17**	-0.40***
	(0.072)	(0.066)
# Employees	-0.0043***	0.0023
	(0.0012)	(0.0014)
$ROA \times Emp$	0.0052***	-0.0031
	(0.0015)	(0.0023)
HHI (Herfindahl Index)	-0.049	-0.0086
	(0.034)	(0.035)
ln R&D expenditures	0.0069	-0.0053
	(0.0073)	(0.021)
ln Advertisement expenditures	-0.00074	-0.0017***
	(0.00092)	(0.00063)
Observations	3,605	3,605

Notes: Robust standard errors clustered at the industry level are shown in parentheses for the last two columns. Variables not shown include firm fixed effects. *** is significant at the 1% level; ** is significant at the 5% level; * is significant at the 10% level.

3.5.2 Peer effects and the decision to engage in CSR.

Even though there is an increasing literature on CSR, little is known about how CSR can influence or be influenced by peer firms or industry dynamics. To my knowledge, only two studies have tried to determine the peer effects of corporate social responsibility (see Liu and Wu 2015; Cao, Liang, and Zhan 2015). Both studies studies present evidence on how firms react to their peers' adoption of CSR. They find that the effects are stronger amongst peers with higher competitive pressure and a more transparent information environment (Cao, Liang, and Zhan, 2015). Furthermore, Liu and Wu (2015) show that when a firm's competitors exhibit a higher level of CSR, the firm is more likely to engage in CSR activities.

A problem that may arise when looking at peer effects, is the probability of a firm's CSR and the competitor's CSR being driven by the same industry level shocks, which could lead to the appearance of peer effects. Similarly, if the firm and its competitors are geographically close, region-level shocks may also affect the interpretation. To deal with this issue, I include industry times year fixed effects and state times year fixed effects in the analysis (see Liu and Wu (2015)). This way, any time-varying trends in industry and region level will be controlled.

Following their results, I find that an increase in the proportion of firms doing CSR within a given industry increases the probability of a firm that was not previously engaged in CSR to decide to adopt it (see Table 3.6). However, this is true up to a point after which the proportion of firms starts to be less relevant in the decision to engage in CSR. The analysis suggests that when 79 percent of the industry engages in CSR, the probability of another firm engaging in CSR given their peers engaging in these activities decreases (see Figure 3.1). Overall, the results coincide with those of Cao, Liang, and Zhan (2015) and Liu and Wu (2015) that a firm's CSR policy is partly shaped by the CSR activities of its competitors. The result is consistent with the theory that firms consider CSR as strategic (Siegel and Vitaliano, 2007; Fernandez-Kranz and Santalo, 2010; Belu and Manescu, 2013) .

The results of the estimations on peer effects on the decision to engage in CSR are provided in Table 3.6, which presents evidence for 3 different dependent variables: CSR (Column (1)), $D-CSR$ (Column (2)), and $Intensity$ (Column (3)). $Intensity$ is measured as the amount of CSR strengths defined by KLD within a firm.[7]

As shown in Table 3.6, the coefficients of $Pfirm$ and $Pfirm^2$ have the expected signs and are significant, indicating that the decision to engage in CSR is affected by their peers in their industry until a point when it becomes less relevant.[8] Column (3) reports how peers in the same industry affect the intensity in which a firm engages in CSR. Contrary to CSR and $D-CSR$, peer effects are not concave for intensity. The higher the percentage of firms engaged in CSR the higher the CSR intensity of a given firm.

So far I have shown that there is evidence that an increase in growth in value added by industry also increases the probability of a firm to engage in Corporate Social Responsibility. This effect seems to be stronger for firms that are not the best performers of their industry. Furthermore, peers within an industry have an influence in the decision of a firm to engage in CSR up to a point. In the following section I also include the analysis of workers attitudes and its effect on the firm's decision to do CSR.

3.5.3 Labor force preferences and CSR.

As explained above, several studies argue that communities' choices put pressure on firms to engage in CSR. Following McWilliams and Siegel (2001), their demand theory suggests that consumers value CSR attributes, and furthermore, companies incorporate CSR into their marketing strategies because "they wish to exploit the appeal of CSR to consumers". However, as Pomering and Dolnicar (2009) suggests,

[7]Intensity is a polytomous and ordinal dependent variable that goes from 0 to 6 if the firm engages in more than one type of CSR: government, diversity, employment, human relations, community, or environment.

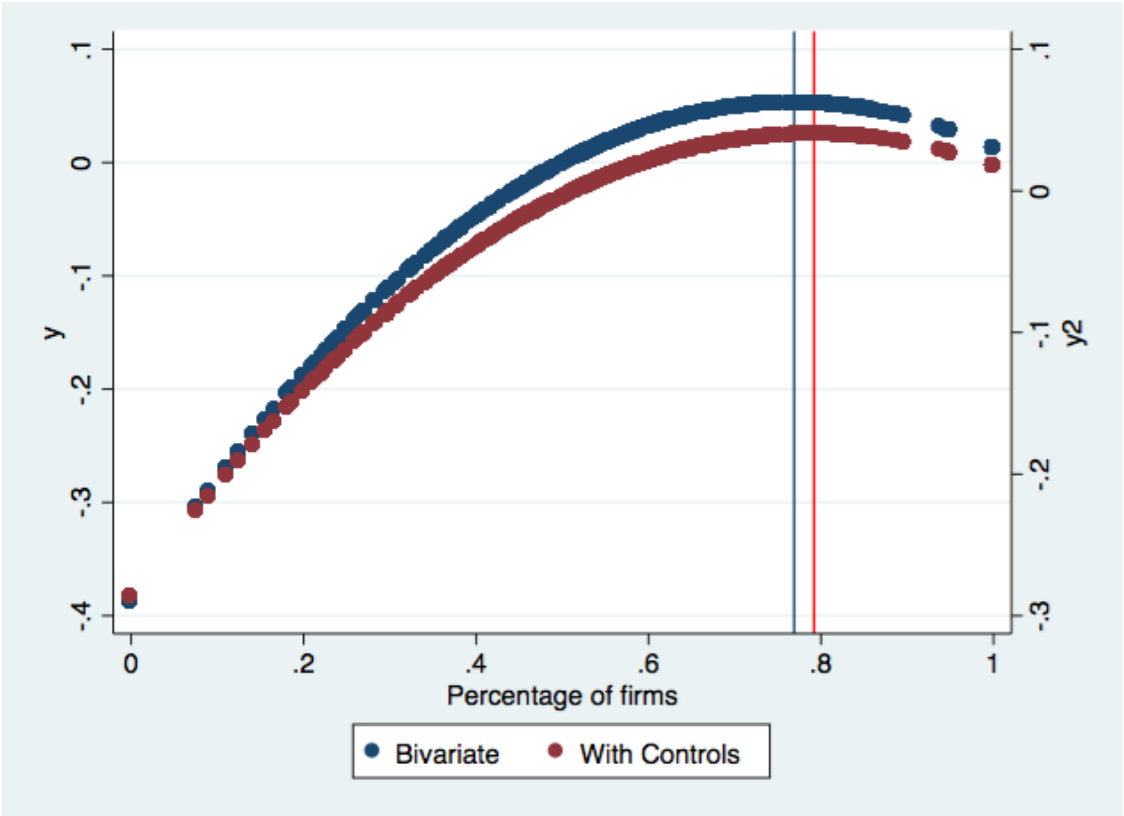
[8]$Pfirm$ and $Pfirm^2$ are jointly significant and different from zero.

Table 3.6: **Peer effects and Corporate Social Responsibility**

	(1) CSR	(2) D-CSR	(3) Intensity
$Pfirms$ (# firms CSR/|# firms Industry)	1.40***	1.35***	1.37***
	(0.10)	(0.12)	(0.31)
$Pfirms^2$	-0.62***	-0.81***	-0.53*
	(0.085)	(0.11)	(0.29)
# Employees	-0.0012	-0.0010	-0.0014
	(0.00078)	(0.00091)	(0.0018)
$ROA \times Emp$	0.00097	0.0011	0.00100
	(0.00081)	(0.00100)	(0.0020)
HHI (Herfindahl Index)	-0.041	-0.10*	0.19
	(0.044)	(0.055)	(0.14)
ln R&D expenditures	0.0021	-0.0019	-0.00072
	(0.0047)	(0.0024)	(0.027)
ln Advertisement expenditures	0.0011	0.0010	-0.0093**
	(0.0015)	(0.0015)	(0.0041)
Firm FE	YES	YES	YES
Time FE	YES	YES	YES
$Geography \times yearFE$	YES	YES	YES
$Industry \times yearFE$	YES	YES	YES
Observations	13,491	13,540	15,743
R-squared	0.084	0.036	0.119
Number of id	2,243	2,250	2,281

Notes: Robust standard errors clustered at the industry level are shown in parentheses. Variables not shown include firm and time fixed effects. *** is significant at the 1% level; ** is significant at the 5% level; * is significant at the 10% level.

Figure 3.1: **Proportion of firms that engage in CSR**



Note: The breaking point for the bivariate regression is 0.769; controlling for firm's characteristics, the breaking point increases to 0.792

worker's awareness levels of CSR activities are low, which proves that the link between CSR and workers is very low. The results in Table 3.7 further confirm this. Overall, worker's attitudes towards specific topics relevant to CSR activities do not have an effect in a firm's decision to engage in these activities.

Table 3.7 refines the sample by only selecting CSR activities that are relevant to community attitudes towards specific topics. Column (1) presents the results for only CSR activities related to diversity. The relationship between attitudes towards LGBT and diversity is positive and significantly different from zero. This suggests that firms consider views towards this attitude relevant enough to engage in diversity CSR. However, attitudes towards unions and the environment have no significant effect on CSR activities.[9] Modeling these specifications with a Logistic regression confirms the result (see results in the Appendix Table 3.B.4).

[9]Unfortunately, given that there is no within-state variation we cannot estimate further results.

The marginal effects point in the same direction with respect to attitudes towards LGBT.

Table 3.7: **LGBT attitudes have a positive effect on a firm's engagement in diversity activities (Diversity-CSR)**

Corporate Social Responsibility			
	(1) Diversity	(2) Unions	(3) Environment
LGBT (CSR diversity index)	0.0017**		
	(0.00079)		
Unions (CSR union relationship index)		-0.10	
		(0.0015)	
Environment (CSR environment index)			0.11
			(0.0012)
# Employees	-0.00059	0.00012	0.00026
	(0.00061)	(0.00098)	(0.00092)
$ROA \times Emp$	0.0014***	0.00036	0.00038
	(0.00039)	(0.00074)	(0.00076)
HHI (Herfindahl Index)	-0.032	-0.058	0.051
	(0.10)	(0.066)	(0.041)
ln R&D expenditures	0.0032	0.010	0.0050
	(0.0061)	(0.0074)	(0.0084)
ln Advertisement expenditures	-0.0026	-0.0038*	0.00070
	(0.0022)	(0.0019)	(0.0018)
Firm FE	YES	YES	YES
Time FE	YES	YES	YES
Observations	9,724	9,724	9,039
R-squared	0.069	0.048	0.058
Number of id	2,141	2,141	2,140

Notes: Robust standard errors clustered at the state level are shown in parentheses. Variables not shown include firm and time fixed effects. *** is significant at the 1% level; ** is significant at the 5% level; * is significant at the 10% level.

3.6 Conclusions

This paper analyses the effect that economic activity, peer pressure and workers preferences have in CSR. The evidence suggests that firms engage in CSR practices in times of economic prosperity for each industry. I also provide further evidence that peer effects are important on the decision of a firm to participate in CSR. When the proportion of firms engaged in CSR increases, the probability of a firm that hasn't engaged before increases. However the proportion of firms engaged in CSR becomes less relevant after almost 80 percent.

With respect to workers preferences, I find that pressure from workers can also explain the decision of a firm to adopt CSR but only for diversity topics. Indeed, I find that the labor force preferences towards LGBT have a strong positive effect on the adoption of CSR measures of diversity. However, there was not enough evidence to see an effect of any other workers preferences having an effect on the firm's decision to engage in CSR activities.

These findings could potentially have policy implications. Policymakers aiming at promoting corporate socially responsible behaviors could stimulate CSR activities in a few firms and this would lead other firms to follow. Similarly, since it appears that community behavior does have an effect on firm's decision on engaging with CSR, policymakers could provide initiatives to raise public awareness on corporate social responsibility. Further research to understand the relationship between workers preferences (and CSR awareness) and firm's engagement in corporate social responsibility is needed.

Appendices

3.A Robustness checks

Table 3.A.1: **Firm's decision to engage in CSR and Industry Growth**

Decision to engage in Corporate Social Responsibility				
	(1)	(2)	(3)	(4)
	OLS	$AR(1)$		$AR(\rho)$
$Growth_{jt-1}$	0.15***	0.21***	0.17***	0.15***
	(0.045)	(0.038)	(0.032)	(0.032)
# Employees	-0.00091	-0.00018	-0.00061	-0.00091
	(0.00084)	(0.0012)	(0.00045)	(0.00084)
$ROA \times Emp$	0.0010	0.00030	0.00072	0.0010
	(0.00090)	(0.0012)	(0.00046)	(0.00092)
HHI (Herfindahl Index)	-0.16**	-0.12**	-0.18***	-0.16**
	(0.081)	(0.054)	(0.048)	(0.054)
ln R&D expenditures	0.00098	0.0071	0.00083	0.00098
	(0.0021)	(0.0075)	(0.0033)	(0.0034)
ln Advertisement expenditures	0.0025	0.0033**	0.0023	0.0025
	(0.0016)	(0.0016)	(0.0015)	(0.0015)
Year	-0.013***		-0.012***	-0.013**
	(0.0023)		(0.0014)	(0.0045)
Firm FE	YES	YES	YES	YES
Observations	13,491	11,248	13,139	13,491
Number of id	2,243	2,144	2,123	
Number of groups				2,243

Notes: Robust standard errors clustered at the industry level are shown in parentheses. Variables not shown include firm fixed effects. *** is significant at the 1% level; ** is significant at the 5% level; * is significant at the 10% level.

Table 3.A.2: **Marginal effects of an increase in positive attitudes towards topics related with CSR**

Strengths of Corporate Social Responsibility			
	(1) Diversity	(2) Unions	(3) Environment
LGBT (CSR diversity index)	0.69***		
	(0.18)		
Unions (CSR union relations index)		0.058	
		(0.14)	
Environment (CSR environment index)			-0.076
			(0.13)
# Employees	-0.00095	0.00012	-0.00051
	(0.00059)	(0.00054)	(0.0018)
$ROA \times Emp$	0.0029***	-0.00048	0.00045
	(0.00097)	(0.00057)	(0.0021)
HHI (Herfindahl Index)	-0.049	-0.076	-0.26***
	(0.047)	(0.053)	(0.076)
ln R&D expenditures	0.029***	0.041**	0.0043
	(0.010)	(0.017)	(0.0074)
ln Advertisement expenditures	0.0058***	0.00049	-0.00090
	(0.0013)	(0.0014)	(0.00087)
Observations	9,724	9,724	9,039

Notes: Robust standard errors clustered at the state level are shown in parentheses. *** is significant at the 1% level; ** is significant at the 5% level; * is significant at the 10% level.

3.B Lagged Regressions

Table 3.B.1: **Lagged: Corporate Social Responsibility and Growth of Value Added by Industry**

Corporate Social Responsibility				
	(1)	(2)	(3)	(4)
	OLS	First Order Autocorrelation		Spatial Autocorrelation
$Growth_{jt-1}$	0.16***	0.095***	0.10***	0.16***
	(0.051)	(0.029)	(0.025)	(0.042)
# Employees	0.078	0.79	0.14	0.078
	(0.57)	(0.99)	(0.43)	(0.41)
$ROA \times Emp$	-0.32	-0.89	-0.18	-0.32
	(0.49)	(0.98)	(0.35)	(0.46)
HHI (Herfindahl Index)	-0.11	-0.13***	-0.19***	-0.11
	(0.092)	(0.050)	(0.059)	(0.069)
ln R&D expenditures	-0.0022	-0.0038	-0.0036	-0.0022
	(0.0038)	(0.0056)	(0.0035)	(0.0042)
ln Advertisement expenditures	0.00055	0.0035**	0.0011	0.00055
	(0.0021)	(0.0018)	(0.0014)	(0.00094)
Year	0.013***		0.0088***	0.013**
	(0.0026)		(0.0014)	(0.0046)
Observations	13,474	11,237	13,132	13,474
Number of id	2,237	2,146	2,126	
Number of groups				2,243

Notes: Robust standard errors clustered at the industry level are shown in parentheses. Variables not shown include firm fixed effects. *** is significant at the 1% level; ** is significant at the 5% level; * is significant at the 10% level.

Table 3.B.2: **Lagged: Heterogeneous treatment effects: Gazelles vs. Elephants**

Corporate Social Responsibility		
	(1) Elephants	(2) Gazelles
$Growth_{jt-1}$	0.22***	-0.0015
	(0.059)	(0.080)
# Employees	0.13	-1.33
	(1.00)	(2.37)
$ROA \times Emp$	-0.027	1.30
	(0.53)	(2.48)
ln R&D expenditures	-0.0080	0.00092
	(0.0054)	(0.0043)
ln Advertisement expenditures	0.0026	-0.0087
	(0.0019)	(0.0052)
Year	0.012***	0.014***
	(0.0032)	(0.0038)
Observations	10,025	3,466
R-squared	0.016	0.020
Number of id	2,032	791

Notes: Robust standard errors clustered at the industry level are shown in parentheses. Variables not shown include firm fixed effects. *** is significant at the 1% level; ** is significant at the 5% level; * is significant at the 10% level.

Table 3.B.3: **Lagged: Multinomial logit marginal effects of the decision to engage in Corporate Social Responsibility and Growth by Industry**

Corporate Social Responsibility		
	(1) Engage	(2) Disengage
$Growth_{jt-1}$	0.19***	-0.40***
	(0.063)	(0.063)
# Employees	-2.45	1.25
	(1.59)	(2.87)
$ROA \times Emp$	2.77	-1.44
	(2.06)	(3.84)
ln R&D expenditures	0.00085	0.011
	(0.0079)	(0.016)
ln Advertisement expenditures	-0.00025	-0.00078
	(0.00095)	(0.00062)
Observations	3,583	3,583

Notes: Robust standard errors clustered at the industry level are shown in parentheses for the last two columns. Variables not shown include firm fixed effects. *** is significant at the 1% level; ** is significant at the 5% level; * is significant at the 10% level.

Table 3.B.4: **Lagged: Marginal effects of an increase in positive attitudes towards topics related with CSR**

Strengths of Corporate Social Responsibility			
	(1) Diversity	(2) Unions	(3) Environment
LGBT (CSR diversity index)	0.76***		
	(0.19)		
Unions (CSR union relations index)		0.12	
		(0.16)	
Environment (CSR environment index)			-0.096
			(0.15)
# Employees	-0.63	-0.18	-1.21
	(0.62)	(0.74)	(2.41)
$ROA \times Emp$	2.17**	-0.24	1.09
	(1.03)	(0.78)	(2.73)
ln R&D expenditures	0.034**	0.043**	0.015
	(0.015)	(0.018)	(0.0096)
ln Advertisement expenditures	0.0064***	0.000087	-0.0013
	(0.0014)	(0.0015)	(0.00094)
Observations	8,039	8,039	7,718

Notes: Robust standard errors clustered at the state level are shown in parentheses. *** is significant at the 1% level; ** is significant at the 5% level; * is significant at the 10% level.

Bibliography

Bagnoli, Mark and Susan G. Watts (2003). "Selling to Socially Responsible Consumers: Competition and the Private Provision of Public Goods". In: *Journal of Economics and Management Strategy* 12.3, pp. 419–445.

Baron, David P. (2001). "Private Politics, Corporate Social Responsibility and Integrated Strategy". In: *Journal of Economics and Management Strategy* 10.1, pp. 7–45.

Belu, Constantin and Cristiana Manescu (2013). "Strategic corporate social responsibility and economic performance". In: *Applied Economics* 45.19, pp. 2751–2764.

Branca, Ana S., Joaquim Pina, and Margarida Catalao-Lopes (2012). "Corporate Giving, Competition and the Economic Cycle". In: *Working Paper* 0874-4548.

Campbell, John L. (2007). "Why Would Corporations Behave in Socially Responsible Ways? An Institutional Theory of Corporate Social Responsibility". In: *Academy of Management Review* 32.3, pp. 946–967.

Cao, Jie, Hao Liang, and Xintong Zhan (2015). *Peer Effects of Corporate Social Responsibility.*

Driscoll, John C. and Aart C. Kraay (1998). "Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data". In: *Review of Economics and Statistics* 40.4, pp. 549–560.

Fernandez-Kranz, Daniel and Jan Santalo (2010). "When Necessity Becomes a Virtue: The Effect of Product Market Competition on Corporate Social Responsibility". In: *Journal of Economics and Management Strategy* 19.2, pp. 453–487.

Galaskiewicz, Joseph and Stanley Wasserman (1989). "Mimetic Processes and Interorganizational Field: An Empirical Test". In: *Administrative Science quarterly* 34.3, pp. 454–479.

Hsiang-Lin, Chih, Chih Hsiang-Hsuan, and Chen Tzu-Yin (2010). "On the Determinants of Corporate Social Responsibility: International Evidence on the Financial Industry". In: *Journal of Business Ethics* 93.1, pp. 115–135.

Hull, Clyde Eirikur and Sandra Rothenberg (2008). "Firm Performance: The interactions of Corporate Social Performance with Innovation and Industry Differentiation". In: *Strategic Management Journal* 29, pp. 781–789.

Jenkins, Rhys (2005). "Globalization, Corporate Social Responsibility and poverty". In: *International Affairs* 81.3, pp. 525–540.

Kitzmueller, Marcus and Jay Shimshack (2012). "Economic Perspectives on Corporate Social Responsibility". In: *Journal of Economic Literature* 50.1, pp. 51–84.

Lankoski, Leena (2009). "Differential Economic Impacts of Corporate Responsibility Issues". In: *Business and Society* 48.2, pp. 206–224.

Liu, Sibo and Dejun Wu (2015). "Competing by conducting good deeds: The peer effect of corporate social responsibility". In: *Finance Research Letters* 1.8.

Lundgren, Tommy (2011). "A Microeconomic Model of Corporate Social Responsibility". In: *Metroeconomica* 62.1, pp. 69–95.

Lyon, Thomas Peyton and John W. Maxwell (2007). *Corporate Social Responsibility and the Environment: A Theoretical Perspective*. Working Papers 2007-16. Indiana University, Kelley School of Business, Department of Business Economics and Public Policy.

Margolis, Joshua D., Hillary Anger Elfenbein, and James P. Walsh (2007). *Does it pay to be good? A Meta-Analysis and redirection of research on the relationship between Corporate Social and Financial Performance*. Working Papers. Harvard Business School.

Margolis, Joshua D and James P Walsh (2003). "Misery Loves Companies: Rethinking Social Initiatives by Business". In: *Administrative science quarterly.* 48.2, p. 268.

Marquis, C., M. A. Glynn, and G. F. Davis (2007). "Community isomorphism and Corporate Social Action". In: *Academy of Management Review* 32.3, pp. 925–945.

Maxwell, John W., Thomas P. Lyon, and Steven C. Hackett (1998). *Self-regulation and social welfare : the political economy of corporate environmentalism*. Milano, Italy: Fondazione Eni Enrico Mattei.

McWilliams, Abagail and Donald S. Siegel (2001). "Corporate Social Responsibility: A Theory of the Firm Perspective". In: *The Academy of Management Review* 26.1, pp. 117–127.

McWilliams, Abagail, Donald S. Siegel, and Patrick M. Wright (2006). "Corporate Social Responsibility: Strategic Implications". In: *Journal of Management Studies* 43.1, pp. 1–18.

Navarro, Peter (1988). "Why do Corporations give to Charity?" In: *The Journal of Business* 61.1, pp. 65–93.

Orlitzky, Marc, Frank L. Schmidt, and Sara L. Rynes (2003). "Corporate Social and Financial Performance: A Meta-Analysis". In: *Organization Studies* 24.3, pp. 403–441.

Pomering, Alan and Sara Dolnicar (2009). "Assessing the prerequisite of successful CSR implementation: are consumers aware of CSR initiatives?" In: *Journal of Business Ethics* 85.2, pp. 285–301.

Reinhardt, Forest L., Robert N. Stavins, and Richard H. K. Vietor (2008). "Corporate Social Responsibility Through an Economic Lens". In: *Review of Environmental Economics and Policy* 2.2, pp. 219–239.

Siegel, Donald S. and Donald F. Vitaliano (2007). "An empirical analysis of the strategic use of Corporate Social Responsibility". In: *Journal of Economics and Management Strategy* 16.3, pp. 773–792.

Ullmann, A. A. (1985). "Data in Search of a Theory: A Critical Examination of the Relationships Among Social Performance, Social Disclosure, and Economic Performance of U.S. Firms." In: *Academy of Management Review Academy of Management Review* 10.3, pp. 540–557.

Wooldridge, Jeffrey M. (2002). *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.

Appendices

3.A Stata code

```
*********************************************
*********************************************/*

Elaborated by Eugenia Suarez as part of my dissertation project
esuare@essex.ac.uk
January 2016

This do-file presents the results of the paper "Informality and
taxes in Mexico"

OUTCOME TABLES:  Table1-Table7.tex
FIGURES: chtaxrate_‘t’.png  for t=2006/2014
chtaxrateall.png (pooling all years)
test_t_t+1.png  for t=2005/2013
taxaveall.png (pooling all years)
SOURCES:
Microdata from INEGI
http://www.inegi.org.mx
A panel of individual may be constructed as 20% of the sample is
followed during 5 quarters
Variable definitions from
(http://www3.inegi.org.mx/rnmpre/index.php/catalog)
*********************************************
*********************************************/
version 13.1
clear all
drop _all
set linesize 255
set mem 5000M
set more off, permanently

run $dod/00_run_first.do
```

```
run $dod/08_aux_dataprep.do
run $dod/08_aux_figures.do

*********************************[]*****************************
********************** SUMMARY STATISTICS TABLES ****************
*********************************************************************
keep if out_new!=1

** Table 1: Summary statistics
** All/Men/Women
sutex age schooly infor incomea grossincomeORIG taxprev_incpr ///
chtaxratepastinc if age>=.16 & age<=.65, minmax
sutex age schooly infor incomea grossincomeORIG taxprev_incpr ///
chtaxratepastinc if age>=.16 & age<=.65 & gender==1, minmax
sutex age schooly infor incomea grossincomeORIG taxprev_incpr ///
chtaxratepastinc if age>=.16 & age<=.65 & gender==2, minmax

** Table 2: Sample Statistics by Labor Market State and Gender
//Men
sutex age schooly infor incomea grossincomeORIG taxprev_incpr ///
chtaxratepastinc if age>=.16 & age<=.65 & gend==1 & infor==1, ///
file($td/desc.tex) replace

sutex age schooly infor incomea grossincomeORIG taxprev_incpr ///
chtaxratepastinc if age>=.16 & age<=.65 & gend==1 & infor==0, ///
file($td/desc.tex) append

//Women
sutex age schooly infor incomea grossincomeORIG taxprev_incpr ///
chtaxratepastinc if age>=.16 & age<=.65 & gend==2 & infor==1, ///
file($td/desc.tex) append

sutex age schooly infor incomea grossincomeORIG taxprev_incpr ///
chtaxratepastinc if age>=.16 & age<=.65 & gend==2 & infor==0, ///
file($td/desc.tex) append

*********************************************************************
******************** Table 1. Main results **********************
*********************************************************************
//All controls
xi: ivreg29 chinfor lL4gross(chtaxratecurri=chtaxratepasti), ///
cluster(entity) first r
outreg2 using $td/table1, tex replace label less(1)

xi: ivreg29 chinfor lL4gross i.scian i.qdate i.city age  ///
agesq (chtaxratecurri=chtaxratepasti), cluster(entity) ///
```

```
partial(i.scian i.qdate i.city ) first r
outreg2 using $td/table1, tex append label less(1)

//OLS all controls
reg chinfor chtaxratepasti lL4gross i.scian i.qdate i.cit  ///
age agesq, r  cluster(entity)
outreg2 using $td/table1, tex append label less(1)

// Women vs. Men
xi: ivreg29 chinfor lL4gross i.scian i.qdate i.city i.marr ///
age agesq (chtaxratecurri=chtaxratepasti) if gend==2, ///
cluster(entity) partial(i.scian i.qdate i.city) first r
outreg2 using $td/table1, tex append label less(1)

reg chinfor chtaxratepasti lL4gross i.scian i.qdate i.city ///
i.married age agesq if gend==2, r  cluster(entity)
outreg2 using $td/table1, tex append label less(1)

xi: ivreg29 chinfor lL4gross i.scian i.qdate i.city age  ///
agesq (chtaxratecurri =chtaxratepasti) if gend==1,  ///
clus(entity) partial(i.scian i.qdate i.city) first r
outreg2 using $td/table1, tex append label less(1)

reg chinfor chtaxratepasti lL4gross i.scian i.qdate i.city  ///
age agesq if gend==1, r  cluster(entity)
outreg2 using $td/table1, tex append label less(1)

*********************************************************************
******************** Table 5. Income results ********************
*********************************************************************

// Poor vs. Rich
xi: ivreg29 chinfor lL4gross i.scian i.qdate i.gend i.city  ///
age agesq (chtaxratecurri=chtaxratepasti) ///
if inc_prevy<62400, ///
cluster(entity) partial(i.scian i.qdate i.gender i.city) first r
outreg2 using $td/table5, tex replace label less(1)

xi: ivreg29 chinfor lL4gross i.scian i.qdate i.gend i.city  ///
age agesq (chtaxratecurri=chtaxratepasti) ///
if inc_prevy>=62400, ///
cluster(entity) partial(i.scian i.qdate i.gender i.city) first r
outreg2 using $td/table5, tex append label less(1)

// Women & Low-Income
xi: ivreg29 chinfor lL4gross i.scian i.qdate i.city i.marr  ///
```

```
age agesq (chtaxratecurri=chtaxratepasti)  ///
if inc_prevy<62400 & gend==2, ///
cluster(entity) partial(i.scian i.qdate i.city ) first r
outreg2 using $td/table5, tex append label less(1)

xi: ivreg29 chinfor lL4gross i.scian i.qdate i.city i.marr ///
age agesq (chtaxratecurri=chtaxratepasti)  ///
if inc_prevy>=62400 & gend==2, ///
cluster(entity) partial(i.scian i.qdate i.city) first r
outreg2 using $td/table5, tex append label less(1)

// Men & Low-Income
xi: ivreg29 chinfor lL4gross i.scian i.qdate i.city  ///
age agesq (chtaxratecurri=chtaxratepasti) ///
if inc_prevy<62400 & gend==1, ///
cluster(entity) partial(i.scian i.qdate i.city ) first r
outreg2 using $td/table5, tex append label less(1)

xi: ivreg29 chinfor lL4gross i.scian i.qdate i.city ///
age agesq (chtaxratecurri=chtaxratepasti) ///
if inc_prevy>=62400 & gend==1, ///
cluster(entity) partial(i.scian i.qdate i.city ) first r
outreg2 using $td/table5, tex append label less(1)

**********************************************************************
******************** Table 2. Women  results ********************
**********************************************************************
// General Women
xi: ivreg29 chinfor lL4gross i.sci i.qdat i.cit i.marr ///
age agesq (chtaxratecurri=chtaxratepasti) if gend==2, ///
cluster(entity) partial(i.scian i.qdate i.city ) first r
outreg2 using $td/table2, tex replace label less(1)

// School years Women
xi: ivreg29 chinfor lL4gross i.sci i.qdat i.cit i.marr ///
age agesq (chtaxratecurri=chtaxratepasti) if schooly<=10  ///
& gend==2, cluster(entity) partial(i.scian i.qdate i.city ) ///
 first r
outreg2 using $td/table2, tex append label less(1)

xi: ivreg29 chinfor lL4gross i.sci i.qdat i.cit i.marr age ///
agesq (chtaxratecurri=chtaxratepasti) if schooly>10 & gend==2, ///
cluster(entity) partial(i.scian i.qdate i.city ) first r
outreg2 using $td/table2, tex append label less(1)

// Age Women
```

```
xi: ivreg29 chinfor lL4gross i.sci i.qdat i.cit i.marr ///
(chtaxratecurri=chtaxratepasti) if age<.38 & gend==2, ///
cluster(entity) partial(i.scian i.qdate i.city) first r
outreg2 using $td/table2, tex append label less(1)

xi: ivreg29 chinfor lL4gross i.sci i.qdat i.cit i.marr ///
(chtaxratecurri=chtaxratepasti) if age>=.38 & gend==2, ///
cluster(entity) partial(i.scian i.qdate i.city) first r
outreg2 using $td/table2, tex append label less(1)

//Urban vs. Rural Women
* Urban
xi: ivreg29 chinfor lL4gross i.sci i.qdat i.cit i.marr age ///
agesq (chtaxratecurri=chtaxratepasti) if gend==2 & city<=43, ///
cluster(entity) partial(i.scian i.qdate i.city ) first r
outreg2 using $td/table2, tex append label less(1)
* Rural
xi: ivreg29 chinfor lL4gross i.sci i.qdat i.cit i.marrage ///
agesq (chtaxratecurri=chtaxratepasti) if gend==2 & city>=81, ///
cluster(entity) partial(i.scian i.qdate i.city ) first r
outreg2 using $td/table2, tex append label less(1)

*********************************************************************
********************** Table 3. Results Low-income **************
*********************************************************************
// General income
xi: ivreg29 chinfor lL4gross i.sci i.qdat i.gend i.cit age ///
agesq (chtaxratecurri=chtaxratepasti) if inc_prevyear<62400, ///
cluster(entity) partial(i.scian i.qdate i.gend i.city) first r
outreg2 using $td/table3, tex replace label less(1)

// School years Women
xi: ivreg29 chinfor lL4gross i.sci i.qdat i.gend i.cit age ///
agesq (chtaxratecurri=chtaxratepasti) if schooly<=10 & ///
inc_prevyear<62400, clu(entity) ///
partial(i.scian i.qdate i.gend i.city) first r
outreg2 using $td/table3, tex append label less(1)

xi: ivreg29 chinfor lL4gross i.sci i.qdat i.gend i.cit age ///
agesq (chtaxratecurri=chtaxratepasti) if schooly>10 & ///
inc_prevyear<62400, clu(entity)///
 partial(i.scian i.qdate i.gend i.city)  first r
outreg2 using $td/table3, tex append label less(1)

// Age Women
xi: ivreg29 chinfor lL4gross i.sci i.qdat i.gend i.cit ///
```

```
(chtaxratecurri=chtaxratepasti) if age<.38 & inc_prevy<62400, ///
cluster(entity) partial(i.scian i.qdate i.gend i.city) first r
outreg2 using $td/table3, tex append label less(1)

xi: ivreg29 chinfor lL4gross i.sci i.qdat i.gend i.cit ///
(chtaxratecurri=chtaxratepasti) if age>=.38 & inc_prevy<62400, ///
cluster(entity) partial(i.scian i.qdate i.gend i.city) first r
outreg2 using $td/table3, tex append label less(1)

//Urban vs. Rural Income
* Urban
xi: ivreg29 chinfor lL4gross i.sci i.cit i.qdat age agesq ///
(chtaxratecurri=chtaxratepasti) if inc_prevy<62400 & city<=43, ///
cluster(entity) partial(i.sci i.cit i.qdat) first r
outreg2 using $td/table3, tex append label less(1)
* Rural
xi: ivreg29 chinfor lL4gross i.sci i.cit i.qdat age agesq ///
(chtaxratecurri=chtaxratepasti) if inc_prevy<62400 & city>=81, ///
cluster(entity) partial(i.sci i.cit i.qdat) first r
outreg2 using $td/table3, tex append label less(1)

*********************************************************************
****************** Table 3. Poor & Women  results **************
*********************************************************************
// General income & Women
xi: ivreg29 chinfor lL4gross i.sci i.cit i.qdat i.marr age ///
agesq (chtaxratecurri=chtaxratepasti) if inc_prevyear<62400 & ///
gend==2, cl(entity) partial(i.sci i.cit i.qdat) first r
outreg2 using $td/table4, tex replace label less(1)

// School years Income & Women
xi: ivreg29 chinfor lL4gross i.sci i.cit i.qdat i.marr age ///
agesq (chtaxratecurri=chtaxratepasti) if schooly<=10 & ///
inc_prevyear<62400 & gend==2, cl(entity)  ///
partial(i.scian i.qdate i.city) first r
outreg2 using $td/table4, tex append label less(1)

xi: ivreg29 chinfor lL4gross i.sci i.cit i.qdat i.marr age ///
agesq (chtaxratecurri=chtaxratepasti) if schooly>10 & ///
inc_prevyear<62400 & gend==2, cl(entity)  ///
partial(i.sci i.cit i.qdat) first r
outreg2 using $td/table4, tex append label less(1)

// Age Women
xi: ivreg29 chinfor lL4gross i.sci i.cit i.qdat i.marr ///
(chtaxratecurri=chtaxratepasti) if age<.38 & inc_prevy<62400 ///
```

```
&  gend==2, cl(entity) partial(i.sci i.cit i.qdat) first r
outreg2 using $td/table4, tex append label less(1)

xi: ivreg29 chinfor lL4gross i.sci i.cit i.qdat i.marr ///
(chtaxratecurri=chtaxratepasti) if age>=.38 & inc_prevy<62400 ///
& gend==2, cl(entity) partial(i.sci i.cit i.qdat) first r
outreg2 using $td/table4, tex append label less(1)

//Urban vs. Rural Income
* Urban
xi: ivreg29 chinfor lL4gross i.sci i.cit i.qdat i.marr age ///
agesq (chtaxratecurri=chtaxratepasti) if inc_prevy<62400 ///
 & city<=43 & gend==2, cl(entity) partial(i.sci i.cit i.qdat) ///
first r
outreg2 using $td/table4, tex append label less(1)
* Rural
xi: ivreg29 chinfor lL4gross i.sci i.cit i.qdat i.marr age ///
agesq (chtaxratecurri=chtaxratepasti) if inc_prevy<62400  ///
& city>=81 & gend==2, cl(entity) partial(i.sci i.cit i.qdat) ///
first r
outreg2 using $td/table4, tex append label less(1)

**********************************************************************
********************** Table 6. Men  results **********************
**********************************************************************
// General Men
xi: ivreg29 chinfor lL4gross i.sci i.cit i.qdat age agesq ///
(chtaxratecurri=chtaxratepasti) if gend==1, cl(entity) ///
partial(i.scian i.qdate i.city) first r
outreg2 using $td/table6, tex replace label less(1)

// School years Men
xi: ivreg29 chinfor lL4gross i.sci i.cit i.qdat age agesq ///
(chtaxratecurri=chtaxratepasti) if schooly<=10 & gend==1, ///
cl(entity) partial(i.scian i.qdate i.city) first r
outreg2 using $td/table6, tex append label less(1)

xi: ivreg29 chinfor lL4gross i.sci i.cit i.qdat age agesq ///
(chtaxratecurri=chtaxratepasti) if schooly>10 & gend==1, ///
cl(entity) partial(i.sci i.cit i.qdat) first r
outreg2 using $td/table6, tex append label less(1)

// Age Men
xi: ivreg29 chinfor lL4gross i.sci i.cit i.qdat ///
(chtaxratecurri=chtaxratepasti) if age<.38 & gend==1, ///
cl(entity) partial(i.sci i.cit i.qdat) first r
```

```
outreg2 using $td/table6, tex append label less(1)

xi: ivreg29 chinfor lL4gross i.sci i.cit i.qdat ///
(chtaxratecurri=chtaxratepasti) if age>=.38 & gend==1, ///
cl(entity) partial(i.sci i.cit i.qdat) first r
outreg2 using $td/table6, tex append label less(1)

//Urban vs. Rural Men
* Urban
xi: ivreg29 chinfor lL4gross i.sci i.cit i.qdat age agesq ///
(chtaxratecurri=chtaxratepasti) if gend==1 & city<=43, ///
cl(entity) partial(i.sci i.cit i.qdat) first r
outreg2 using $td/table6, tex append label less(1)
* Rural
xi: ivreg29 chinfor lL4gross i.sci i.cit i.qdat age agesq ///
(chtaxratecurri=chtaxratepasti) if gend==1 & city>=81, ///
cl(entity) partial(i.sci i.cit i.qdat) first r
outreg2 using $td/table6, tex append label less(1)

*********************************************************************
*************** Table 7. Men  & Low-Income results *************
*********************************************************************
// General income & Men
xi: ivreg29 chinfor lL4gross i.sci i.cit i.qdat age agesq ///
(chtaxratecurri=chtaxratepasti) if inc_prevy<62400 & gend==1, ///
cl(entity) partial(i.sci i.cit i.qdat) first r
outreg2 using $td/table7, tex replace label less(1)

// School years Income & Men
xi: ivreg29 chinfor lL4gross i.sci i.cit i.qdat age agesq ///
(chtaxratecurri=chtaxratepasti) if schooly<=10 & inc_prevy<62400 ///
&  gend==1, cl(entity) partial(i.sci i.cit i.qdat) first r
outreg2 using $td/table7, tex append label less(1)

xi: ivreg29 chinfor lL4gross i.sci i.cit i.qdat age agesq ///
(chtaxratecurri=chtaxratepasti) if schooly>10 & inc_prevy<62400 ///
& gend==1, cl(entity) partial(i.scian i.qdate i.city ) first r
outreg2 using $td/table7, tex append label less(1)

// Age Men
xi: ivreg29 chinfor lL4gross i.sci i.cit i.qdat ///
(chtaxratecurri=chtaxratepasti) if age<.38 & inc_prevy<62400 ///
& gend==1, cl(entity) partial(i.sci i.cit i.qdat) first r
outreg2 using $td/table7, tex append label less(1)

xi: ivreg29 chinfor lL4gross i.sci i.cit i.qdat ///
```

```
(chtaxratecurri=chtaxratepasti) if age>=.38 & inc_prevy<62400 ///
& gend==1, cl(entity) partial(i.sci i.cit i.qdat) first r
outreg2 using $td/table7, tex append label less(1)

//Urban vs. Rural Income Men
* Urban
xi: ivreg29 chinfor lL4gross i.sci i.cit i.qdat age agesq ///
(chtaxratecurri=chtaxratepasti) if inc_prevy<62400 & city<=43 ///
& gend==1, cl(entity) partial(i.sci i.cit i.qdat) first r
outreg2 using $td/table7, tex append label less(1)
* Rural
xi: ivreg29 chinfor lL4gross i.sci i.cit i.qdat age agesq ///
(chtaxratecurri=chtaxratepasti) if inc_prevy<62400 & city>=81 ///
& gend==1, cl(entity) partial(i.sci i.cit i.qdat) first r
outreg2 using $td/table7, tex append label less(1)


**********************************************
**********************************************/*

Elaborated by Eugenia Suarez as part of my dissertation project
esuare@essex.ac.uk
January 2016

This do-file presents the results of the paper "Latin American
Aid Patterns:Democratic Political Regimes and US Domestic
Politics"

OUTCOME TABLES:  Table1-Table5.tex  (Economic Aid)
Table1a-Table5a.tex (Military Aid)
Table6  (Heckman)
SOURCES:
Bueno de Mesquita, 2007
Tingley, 2010
Easterly, 2013
COWTRADE (http://cow.la.psu.edu/COW2%20Data/Trade/Trade.html)
**********************************************
**********************************************/
version 13.1
clear all
drop _all
set linesize 255
set mem 5000M
set more off, permanently

run $dod/00_run_first.do
```

```
use $dd/Complete_data.dta, clear

macro drop cont
global cont "eS RB RB2 lnPOP tau LlnTRADE x1 x2 x3 x4 x5 Lciv"
global outr "label nocons obs td(3) less(0) bdec(3)"
global clu "cluster(ccode) robust"
*********************************************************************
****************************** Estimation ***************************
*********************************************************************

*********************************************************************
******************Table 1: Main specification
***** Amount of aid received
** Column 1
xtreg lnEaid  W $cont RUSinfluence i.year if ///
region!="LAC", fe $clu
outreg2 using "$td/table1",tex replace $outr
** Column 2
xtreg lnEaid  W $cont RUSinfluence i.year if ///
region=="LAC", fe $clu
outreg2 using "$td/table1", tex append $outr
** Column 3
xtreg lnEaid USinfluence W $cont RUSinfluence i.year if ///
region!="LAC", fe $clu
outreg2 using "$td/table1",tex append $outr
** Column 4
xtreg lnEaid USinfluence W $cont RUSinfluence i.year if ///
region=="LAC", fe $clu
outreg2 using "$td/table1", tex append $outr

***** Probability of receiving aid
** Column 5
xtreg Eaid  W $cont RUSinfluence i.year if ///
region!="LAC", fe $clu
*margins, dydx (*) post
outreg2 using "$td/table1", tex append $outr
** Column 6
xtreg Eaid  W $cont  RUSinfluence i.year if ///
region=="LAC", fe $clu
outreg2 using "$td/table1", tex append $outr
** Column 7
xtreg Eaid USinfluence W $cont RUSinfluence i.year if ///
region!="LAC", fe $clu
*margins, dydx (*) post
outreg2 using "$td/table1", tex append $outr
** Column 8
```

```
xtreg Eaid USinfluence W $cont  RUSinfluence i.year if ///
region=="LAC", fe $clu
outreg2 using "$td/table1", tex append $outr

*****************************************************************
******************Table 2: Lciv LConInd lnPOP tau (table 4)
*** REST OF THE WORLD
** Column 1
xtreg lnEaid USinfluence W ideo wideo wint $cont RUSinfluence ///
new_leader_dum year_in_office  i.year  if region!="LAC" , ///
fe clus(ccode) robust
outreg2 using "$td/table2", tex replace $outr
*** LAC
** Column 2
xtreg lnEaid USinfluence W ideo wideo wint $cont RUSinfluence ///
new_leader_dum year_in_office i.year if region=="LAC", fe $clus
outreg2 using "$td/table2", tex append $outr

***** Probability of receiving aid
*** REST OF THE WORLD
** Column 3
xtreg Eaid USinfluence W ideo wideo wint $cont if region!="LAC",fe ///
$clus
outreg2 using "$td/table2", tex append $outr
*** LAC
** Column 4
xtreg Eaid USinfluence W ideo wideo wint $cont if region=="LAC",fe ///
$cluse
outreg2 using "$td/table2", tex append $outr

*****************************************************************
******************Table 3: Descriptive Statistics
** Table 3: Summary statistics
sutex lnEaid Eaid lnMaid Maid USinfluence W ideology wideo ///
wint eS RB ln_per_capita_income if region!="LAC"
sutex lnEaid Eaid lnMaid Maid USinfluence W ideology wideo ///
wint eS RB ln_per_capita_income if region=="LAC"

*****************************************************************
******************Table 4: Including US exports
*** REST OF THE WORLD
** Column 1
xtreg lnEaid USinfluence W ideo wideo wint usexpor $cont ///
new_leader_dum year_in_office i.year if region!="LAC",fe $clus
outreg2 using "$td/table4", tex replace $outr
*** LAC
```

```
** Column 2
xtreg lnEaid USinfluence W ideo wideo wint usexpor $cont ///
new_leader_dum year_in_office i.year if region=="LAC",fe $clus
outreg2 using "$td/table4", tex append $outr

***** Probability of receiving aid
*** REST OF THE WORLD
** Column 3
xtreg Eaid USinfluence  W ideo wideo wint usexports $cont ///
new_leader_dum year_in_office if region!="LAC", fe $clus
outreg2 using "$td/table4", tex append $outr
*** LAC
** Column 4
xtreg Eaid USinfluence  W ideo wideo wint usexports $cont ///
new_leader_dum year_in_office if region=="LAC", fe $clus
outreg2 using "$td/table4", tex append $outr

****************************************************************
******************Table 5:Guerrillas & Civil War & Reagan
*** REST OF THE WORLD
** Column 1
xtreg lnEaid USinfluence W ideo wideo wint Guerilla reagan $cont ///
new_leader_dum year_in_office i.year if region!="LAC", fe $clus
outreg2 using "$td/table5", tex replace $outr
*** LAC
** Column 2
xtreg lnEaid USinfluence W ideo wideo wint Guerilla reagan $cont ///
new_leader_dum year_in_office i.year if region=="LAC", fe $clus
outreg2 using "$td/table5", tex append $outr

***** Probability of receiving aid
*** REST OF THE WORLD
** Column 3
xtreg Eaid USinfluence W ideo wideo wint Guerilla reagan $cont ///
new_leader_dum year_in_office if region!="LAC", fe $clus
outreg2 using "$td/table5", tex append $outr
*** LAC
** Column 4
xtreg Eaid USinfluence W ideo wideo wint Guerilla reagan $cont ///
new_leader_dum year_in_office if region=="LAC", fe $clus
outreg2 using "$td/table5", tex append $outr

****************************************************************
****************************** Appendix Tables *****************
******************************  MILITARY AID  ******************
****************************************************************
```

```
******************Table 1: Main specification
***** Amount of aid received
** Column 1
xtreg lnMaid USinfluence W $cont i.year if region!="LAC",fe $clus
outreg2 using "$td/table1a", tex replace $outr
** Column 2
xtreg lnMaid USinfluence W $cont i.year if region=="LAC",fe $clus
outreg2 using "$td/table1a", tex append $outr

***** Probability of receiving aid
** Column 3
xtreg Maid USinfluence W $cont i.year if region!="LAC",fe $clus
outreg2 using "$td/table1a", tex append $outr
** Column 4
xtreg Maid USinfluence W $cont i.year if region=="LAC", fe $clus
outreg2 using "$td/table1a", tex append $outr

*****************************************************************
******************Table 2:  Lciv LConInd lnPOP tau (table 4)
*** REST OF THE WORLD
** Column 1
xtreg lnMaid USinfluence W wideo wint $cont i.year if ///
region!="LAC", fe $clus
outreg2 using "$td/table2a", tex replace $outr
*** LAC
** Column 2
xtreg lnMaid USinfluence W wideo wint $cont i.year if ///
region=="LAC", fe $clus
outreg2 using "$td/table2a", tex append  $outr

***** Probability of receiving aid
*** REST OF THE WORLD
** Column 3
xtreg Maid USinfluence W wideo wint $cont i.year if ///
region!="LAC", fe $clus
outreg2 using "$td/table2a", tex append $outr
*** LAC
** Column 4
xtreg Maid USinfluence W wideo wint $cont i.year if ///
region=="LAC", fe $clus
outreg2 using "$td/table2a", tex append $outr

*****************************************************************
******************Table 4: Including US exports
*** REST OF THE WORLD
** Column 1
```

```
xtreg lnMaid USinfluence W wideo wint usexpor $cont i.year ///
if region!="LAC", fe $clus
outreg2 using "$td/table4a", tex replace $outr
*** LAC
** Column 2
xtreg lnMaid USinfluence W wideo wint usexpor $cont i.year ///
if region=="LAC", fe $clus
outreg2 using "$td/table4a", tex append $outr

***** Probability of receiving aid
*** REST OF THE WORLD
** Column 3
xtreg Maid USinfluence W wideo wint usexpor $cont i.year ///
if region!="LAC", fe $clus
outreg2 using "$td/table4a", tex append $outr
*** LAC
** Column 4
xtreg Maid USinfluence W wideo wint usexpor $cont i.year ///
if region=="LAC", fe $clus
outreg2 using "$td/table4a", tex append $outr

*********************************************************************
******************Table 5:  Guerrillas & Civil War & Reagan
*** REST OF THE WORLD
** Column 1
xtreg lnMaid USinfluence W wideo wint Guerilla reagan ///
$cont i.year if region!="LAC", fe $clus
outreg2 using "$td/table5a", tex replace $outr
*** LAC
** Column 2
xtreg lnMaid USinfluence W wideo wint Guerilla reagan ///
$cont i.year if region=="LAC", fe $clus
outreg2 using "$td/table5a", tex append $outr

***** Probability of receiving aid
*** REST OF THE WORLD
** Column 3
xtreg Maid USinfluence W wideo wint Guerilla reagan ///
$cont i.year if region!="LAC", fe $clus
outreg2 using "$td/table5a", tex append $outr
*** LAC
** Column 4
xtreg Maid USinfluence W wideo wint Guerilla reagan ///
$cont i.year if region=="LAC", fe $clus
outreg2 using "$td/table5a", tex append $outr
```

```
fvset base 2 cty
global cont1 "eS RB RB2 lnPOP tau LlnTRADE Lciv"

*** Heckman Sample Selection Model
heckman lnEaid USinfluence W $cont1 wideo wint x1 x2 x3 ///
x4 x5 i.cty i.year if region!="LAC", select(Eaid = ///
USinfluence $cont1 W wideo wint USworld RA x1 x2 x3 x4 ///
x5 i.cty i.year) two
outreg2 using "$td/table6", tex replace $outr

heckman lnEaid USinfluence W $cont1 wideo wint x1 x2 x3 ///
x4 x5 i.cty i.year if region=="LAC", select(Eaid = ///
USinfluence $cont1 W wideo wint USworld RA x1 x2 x3 x4 ///
x5 i.cty i.year) two
outreg2 using "$td/table6", tex append $outr

*************************************************************
**** Newey-West robust standard errors (30 lags)

** Main specification
xi: newey2 lnEaid USinfluence W wideo wint $cont ///
RUSinfluence new_leader_dum year_in_office i.cty if ///
region!="LAC", lag(30) i(obs) t(year) force
outreg2 using "$td/newey_bsc",tex replace $outr

xi: newey2 lnEaid USinfluence W wideo wint $cont ///
RUSinfluence new_leader_dum year_in_office i.cty if ///
region=="LAC", lag(30) i(obs) t(year) force
outreg2 using "$td/newey_bsc",tex append $outr

** Including US exports
xi: newey2 lnEaid USinfluence W wideo wint usexpor $cont ///
RUSinfluence new_leader_dum year_in_office i.cty if ///
region!="LAC", lag(30) i(obs) t(year) force
outreg2 using "$td/newey_bsc",tex append $outr

xi: newey2 lnEaid USinfluence W wideo wint usexpor $cont ///
RUSinfluence new_leader_dum year_in_office i.cty if ///
region=="LAC", lag(30) i(obs) t(year) force
outreg2 using "$td/newey_bsc",tex append l$outr




***********************************************
```

```
**********************************************/*

Elaborated by Eugenia Suarez as part of my dissertation project
esuare@essex.ac.uk
January 2016

This do-file presents the results of the paper "CSR and
the Economy"

OUTCOME TABLES:  Table1-Table6.tex
Table8.tex & tablelog.tex (attitudes)
FIGURES: concave.png   (concavity)

SOURCES:
KLD
COMPUSTAT
Both available at: https://wrds-web.wharton.upenn.edu/wrds/)
ANES (http://www.electionstudies.org/)
***********************************************
***********************************************/
version 13.1
clear all
drop _all
set linesize 255
set mem 5000M
set more off, permanently

do "/Users/eugenia/Dropbox/PHD1/CSR/stata/do2/00_run_first.do"
set matsize 3000


use $dd/csr.dta, clear
set more off, perm
xtdes

macro drop control
global control "emp roaemp  hhi lnrd lnxad"
global control_lag "lemp lroaemp  llnrd llnxad"
global cli "cluster(industry)"
global cls "cluster(state)"
sutex act ebit roa emp roaemp hhi lnrd lnxad laggdp_ind ///
mgays munions menviron mbusiness pfirmscsr

xtserial csr laggdp_ind $control

*****************************************************************
```

```
****************** Table 2. CSR and Growth results ************
*************************************************************
**** Column 1:
xtreg csr laggdp_ind $control year, fe  robust $cli
outreg2 using $td/table2, tex replace nolabel less(1)

**** Column 2:
xtlogit csr laggdp_ind $control year, fe
outreg2 using $td/table2, tex append nolabel less(1)

**** Column 3:
xtregar csr laggdp_ind $control, fe
outreg2 using $td/table2, tex append nolabel less(1)

**** Column 4:
xtARGLS csr laggdp_ind $control year if maxt>3 & maxNind>3, ///
groupvar(id) timevar(year) ar(1) ttrendvars(st_16 st_36) $cli
outreg2 using $td/table2, tex append nolabel less(1)

**** Column 5:
xtscc csr laggdp_ind $control year, fe
outreg2 using $td/table2, tex append nolabel less(1)

*************************************************************
*************** Table 3. Gazelles vs Elephants ***************
*************************************************************
** Heterogeneous effects
sum roa, det
gen gazelle=.
replace gazelle=0 if roa<r(p75)
replace gazelle=1 if roa>=r(p75) & roa!=.

**** Column 1: Elephant
xtreg csr laggdp_ind $control year if gazelle==0, fe robust $cli
outreg2 using $td/table3, tex replace nolabel less(1)

**** Column 2: Gazelle
xtreg csr laggdp_ind $control year if gazelle==1, fe robust $cli
outreg2 using $td/table3, tex append nolabel less(1)

*************************************************************
***************** Table 4. Dcsr Mlogit ***********************
*************************************************************

**** Column 1 & 2: Mlogit with cluster by industry
mlogit Dcsr laggdp_ind $control if past==1, cluster(industry)
```

```
est store mlogitcluster
margins, dydx(*) predict(outcome(1))atmeans post
est store eq3
outreg2 using $td/table4, tex replace nolabel less(1)
est restore mlogitcluster
margins, dydx(*) predict(outcome(-1))atmeans post
est store eq4
outreg2 using $td/table4, tex append nolabel less(1)


*****************************************************************
********************* Table for Annex ***************************
*****************************************************************
** Table 5:

**** Column 1: FE without AR process
xtreg Dcsr laggdp_ind $control year, fe robust $cli
outreg2 using $td/table5, tex replace nolabel less(1)

**** Column 2: FE without AR process and with sector dummies
xtregar Dcsr laggdp_ind $control, fe
outreg2 using $td/table5, tex append nolabel less(1)

**** Column 3: FE with AR (1) and cluster by sector
xtARGLS Dcsr laggdp_ind $control year if maxt>3 & maxNind>3, ///
groupvar(id) timevar(year) ar(1) ttrendvars(st_16 st_36) $cli
outreg2 using $td/table5, tex append nolabel less(1)

**** Column 4: FE allowing spatial correlation
xtscc Dcsr laggdp_ind $control year, fe
outreg2 using $td/table5, tex append nolabel less(1)

*****************************************************************
************************ Peer effects ***************************
*****************************************************************
** Graph

*Aux regressions for graph (get breaking points)
xtreg Dcsr pfirmscsr pf2
xtreg Dcsr pfirmscsr pf2 $control, robust $cli

preserve
* For graph use concave.dta and
use $dd/Concave2.dta, clear
twoway ||scatter y x, xline(0.769, lcol(navy)) ///
xtitle("Percentage of firms") yaxis(1)|| ||scatter y2 x2, ///
```

```
xline(0.792, lcolor(red))  yaxis(2)||, legend(rows(1) ///
lab(1 "Bivariate") lab(2 "With Controls"))
graph export "$gd/Concave2.png", replace
restore

merge m:m year industry using $dd/csr.dta
** Table 6:
** Column 1
xi: xtreg csr  pfirmscsr pf2 laggdp_ind $control ran, fe robust $cli
outreg2 using $td/table6, tex replace nolabel less(1)

** Column 2
xtreg  Dcsr  pfirmscsr pf2 $control ran, fe robust $cli
outreg2 using $td/table6, tex append nolabel less(1)

** Column 3
xtreg intensity pfirmscsr pf2 $control  ran, fe robust $cli
outreg2 using $td/table6, tex append nolabel less(1)


**********************************************************************
**********************  Consumer attitudes ***********************
**********************************************************************

/* Table 8: Probability that consumer attitudes could affect
the decision of doing CSR (LPM) */

** Column 1
xtreg div_str mgays $control i.year,  fe robust $cls
outreg2 using $td/table8, tex replace nolabel less(1)
** Column 2
xtreg emp_str munions $control i.year, fe  robust $cls
outreg2 using $td/table8, tex append nolabel less(1)
** Column 3
xtreg env_str menviron $control i.year, fe robust $cls
outreg2 using $td/table8, tex append nolabel less(1)

/* Table log: Probability that consumer attitudes could affect
the decision of doing CSR (logit) */

logit div_str  mg $control i.year, robust $cls
margins, dydx(*) post
outreg2 using $td/tablelog, tex replace nolabel less(1)

logit emp_str  mu $control i.year, robust $cls
margins, dydx(*) post
```

```
outreg2 using $td/tablelog, tex append nolabel less(1)

logit env_str  me $control i.year, robust $cls
margins, dydx(*) post
outreg2 using $td/tablelog, tex append nolabel less(1)
```